PHARMACEUTICAL COMPOSITION CONTAINING A STABILISED mRNA OPTIMISED FOR TRANSLATION IN ITS CODING REGIONS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application is a Continuation-In-Part of co-pending PCT Application No. PCT/EP02/06180 filed June 5, 2002, which in turn, claims priority from German Application No. DE 101 27 283.9, filed June 5, 2001. Applicants claim the benefits of 35 U.S.C. §120 as to the PCT application and priority under 35 U.S.C. §119 as to the said German application, and the disclosures of both applications are incorporated herein in their entireties.

BACKGROUND OF THE INVENTION

Field Of the Invention

[0002] The present invention relates to a pharmaceutical composition containing an mRNA that is stabilised by sequence modifications in the translated region and is optimised for translation. The pharmaceutical composition according to the invention is suitable in particular as an inoculating agent and also as a therapeutic agent for tissue regeneration. Furthermore, a process for determining sequence modifications that stabilise mRNA and optimise mRNA translation is disclosed.

Description of the Prior Art

[0003] Gene therapy and genetic vaccination are tools of molecular medicine whose use in the treatment and prevention of diseases has considerable potential. Both of these approaches are based on the incorporation of nucleic acids into a patient's cells or tissue as well as on the subsequent processing of the information encoded by the incorporated nucleic acids, i.e. the expression of the desired polypeptides.

[0004] Conventional procedures involved in previous applications of gene therapy and genetic vaccination involved the use of DNA in order to incorporate the required genetic information into a cell. In this connection various processes for the incorporation of DNA into cells have been developed, such as for example calcium phosphate transfection, polyprene transfection, protoplast fusion, electroporation, microinjection

and lipofection, in which connection lipofection in particular has proved to be a suitable process.

[0005] A further process that has been suggested in particular for the case of genetic vaccination involves the use of DNA viruses as DNA vehicles. Because such viruses are infectious, a very high transfection rate can be achieved when using DNA viruses as vehicles. The viruses used are genetically altered so that no functional infectious particles are formed in the transfected cell. Despite this precautionary measure, however, the risk of uncontrolled propagation of the introduced therapeutic gene as well as viral genes remains due to the possibility of recombination events.

[0006] Normally DNA incorporated into a cell is integrated to a certain extent into the genome of the transfected cell. On the one hand this phenomenon can exert a desirable effect, since in this way a long-lasting action of the introduced DNA can be achieved. On the other hand the integration into the genome brings with it a significant risk for gene therapy. Such integration events may, for example, involve an insertion of the incorporated DNA into an intact gene, which produces a mutation that interferes with or completely ablates the function of the endogenous gene. As a result of such integration events, enzyme systems that are important for cellular viability may be switched off. Alternatively, there is also the risk of inducing transformation of the transfected cell if the integration site occurs in a gene that is critical for regulating cell growth. Accordingly, when using DNA viruses as therapeutic agents and vaccines, a carcinogenic risk cannot be excluded. In this connection it should also be borne in mind that, in order to achieve effective expression of the genes incorporated into the cell, the corresponding DNA vehicles comprise a strong promoter, for example the viral CMV promoter. The integration of such promoters into the genome of the treated cell may, however, lead to undesirable changes in the regulation of the gene expression in the cell.

[0007] A further disadvantage of the use of DNA as a therapeutic agent or vaccine is the induction of pathogenic anti-DNA antibodies in the patient, resulting in a potentially fatal immune response.

[0008] In contrast to DNA, the use of RNA as a therapeutic agent or vaccine is regarded as significantly safer. In particular, use of RNA is not associated with a risk of stable

integration into the genome of the transfected cell. In addition, no viral sequences such as promoters are necessary for effective transcription of RNA. Beyond this, RNA is degraded rapidly in *vivo*. Indeed, the relatively short half-life of RNA in circulating blood, as compared to that of DNA, reduces the risks associated with developing pathogenic anti-RNA antibodies. Indeed, anti-RNA antibodies have not been detected to date. For these reasons RNA may be regarded as the molecule of choice for molecular medicine therapeutic applications.

[0009] However, some basic problems still have to be solved before medical applications based on RNA expression systems can be widely employed. One of the problems in the use of RNA is the reliable, cell-specific and tissue-specific efficient transfer of the nucleic acid. Since RNA is normally found to be very unstable in solution, up to now RNA could not be used or used only very inefficiently as a therapeutic agent or inoculating agent in the conventional applications designed for DNA use.

[0010] Enzymes that break down RNA, so-called RNases (ribonucleases), are responsible in part for the instability. Even minute contamination by ribonucleases is sufficient to degrade down RNA completely in solution. Moreover, the natural decomposition of mRNA in the cytoplasm of cells is exquisitely regulated. Several mechanisms are known which contribute to this regulation. The terminal structure of a functional mRNA, for example, is of decisive importance. The so-called "cap structure" (a modified guanosine nucleotide) is located at the 5' end and a sequence of up to 200 adenosine nucleotides (the so-called poly-A tail) is located at the 3′ end. The RNA is recognised as mRNA by virtue of these structures and these structures contribute to the regulatory machinery controlling mRNA degradation. In addition there are further mechanisms that stabilise or destabilise RNA. Many of these mechanisims are still unknown, although often an interaction between the RNA and proteins appears to be important in this regard. For example, an mRNA surveillance system has been described (Hellerin and Parker, Annu. Rev. Genet. 1999, 33: 229 to 260), in which incomplete or nonsense mRNA is recognised by specific feedback protein interactions in the cytosol and is made accessible to decomposition. Exonucleases appear to contribute in large measure to this process.

[0011] Certain measures have been proposed in the prior art to improve the stability of RNA and thereby enable its use as a therapeutic agent or RNA vaccine.

[0012] In EP-A-1083232 a process for the incorporation of RNA, in particular mRNA, into cells and organisms has been proposed in order to solve the aforementioned problem of the instability of RNA ex vivo. As described therein, the RNA is present in the form of a complex with a cationic peptide or protein.

[0013] WO 99/14346 describes further processes for stabilising mRNA. In particular, modifications of the mRNA are proposed that stabilise the mRNA species against decomposition by RNases. Such modifications may involve stabilisation by sequence modifications, in particular reduction of the C content and/or U content by base elimination or base substitution. Alternatively, chemical modifications may be used, in particular the use of nucleotide analogues, as well as 5′ and 3′ blocking groups, an increased length of the poly-A tail as well as the complexing of the mRNA with stabilising agents, and combinations of the aforementioned measures.

[0014] In US patents US 5,580,859 and US 6,214,804 mRNA vaccines and mRNA therapeutic agents are disclosed *inter alia* within the scope of "transient gene therapy" (TGT). Various measures are described therein for enhancing the translation efficiency and mRNA stability that relate in particular to the composition of the non-translated sequence regions.

[0015] Bieler and Wagner (in: Schleef (Ed.), Plasmids for Therapy and Vaccination, Chapter 9, pp. 147 to 168, Wiley-VCH, Weinheim, 2001) report on the use of synthetic genes in combination with gene therapy methods employing DNA vaccines and lentiviral vectors. The construction of a synthetic *gag*-gene derived from HIV-1 is described, in which the codons have been modified with respect to the wild type sequence (alternative codon usage) in such a way as to correspond to frequently used codons found in highly expressed mammalian genes. In this way, in particular, the A/T content compared to the wild type sequence was reduced. Moreover, the authors found an increased rate of expression of the synthetic *gag* gene in transfected cells. Furthermore, increased antibody formation against the *gag* protein was observed in mice immunised with the

synthetic DNA construct. An increase in cytokine release in vitro from transfected spleen cells of such mice was also observed. Finally, an induction of a cytotoxic immune response in mice immunised with the *gag* expression plasmid was also found. The authors of this article attribute the improved properties of their DNA vaccine to a change in the nucleocytoplasmic transport of the mRNA expressed by the DNA vaccine, which was due to the optimised codon usage. The authors maintain that the effect of the altered codon usage on the translation efficiency was only slight.

SUMMARY OF THE INVENTION

[0016] The object of the present invention is to provide a new system for gene therapy and genetic vaccination that overcomes the disadvantages associated with the properties of DNA therapeutic agents and DNA vaccines and increases the effectiveness of therapeutic agents based on RNA species.

[0017] This object is achieved by the embodiments of the present invention characterised in the claims.

[0018] In particular, a modified mRNA, as well as a pharmaceutical composition comprising at least one modified mRNA of the present invention and a pharmaceutically compatible carrier and/or vehicle are provided. The modified mRNA encodes at least one biologically active or antigenic peptide or polypeptide, wherein the sequence of the mRNA comprises at least one modification as set forth herein below as compared to the wild type mRNA. Such modifications may be located in the region coding for the at least one peptide or polypeptide, or in untranslated regions.

[0019] In one aspect, the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased relative to that of the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide. The encoded amino acid sequence, however, remains unchanged compared to the wild type. (i.e. silent with respect to the encoded amino acid sequence).

[0020] This modification is based on the fact that, for efficient translation of an mRNA, the sequence of the region of the mRNA to be translated is essential. In this connection the composition and the sequence of the various nucleotides play an important role. In

particular sequences with an increased G (guanosine)/C(cytosine) content are more stable than sequences with an increased A(adenosine)/U (uracil) content. In accordance with the invention, the codons are varied compared to the wild type mRNA, while maintaining the translated amino acid sequence, so that they contain increased amounts of G/C nucleotides. Since several different codons can encode the same amino acid, due to degeneracy of the genetic code, the codons most favourable for the stability of the modified mRNA can be determined and incorporated (alternative codon usage).

[0021] Depending on the amino acid encoded by the modified mRNA, various possibilities for modifying the mRNA sequence compared to the wild type sequence are feasible. In the case of amino acids that are encoded by codons that contain exclusively G or C nucleotides, no modification of the codon is necessary. Thus, the codons for Pro (CCC or CCG), Arg (CGC or CGG), Ala (GCC or GCG) and Gly (GGC or GGG) do not require any alteration since no A or U is present.

[0022] In the following cases the codons that contain A and/or U nucleotides are altered by substituting other codons that code for the same amino acids, but do not contain A and/or U. Examples include: the codons for Pro, which may be changed from CCU or CCA to CCC or CCG; the codons for Arg, which may be changed from CGU or CGA or AGA or AGG to CGC or CGG; the codons for Ala, which may be changed from GCU or GCA to GCC or GCG; the codons for Gly, which may be changed from GGU or GGA to GGC or GGG.

[0023] In other cases, wherein A and/or U nucleotides may not be eliminated from the codons, it is however possible to reduce the A and U content by using codons that contain fewer A and/or U nucleotides. For example: the codons for Phe, which may be changed from UUU to UUC; the codons for Leu, which may be changed from UUA, CUU or CUA to CUC or CUG; the codons for Ser, which may be changed from UCU or UCA or AGU to UCC, UCG or AGC; the codon for Tyr, which may be changed from UAU to UAC; the stop codon UAA, which may be changed to UAG or UGA; the codon for Cys, which may be changed from UGU to UGC; the codon for His, which may be changed from CAU to CAC; the codon for Gln, which may be changed from CAA to CAG; the codons for Ile, which may be changed from AUU or AUA to AUC;

the codons for Thr, which may be changed from ACU or ACA to ACC or ACG; the codon for Asn, which may be changed from AAU to AAC; the codon for Lys, which may be changed from AAA to AAG; the codons for Val, which may be changed from GUU or GUA to GUC or GUG; the codon for Asp, which may be changed from GAU to GAC; the codon for Glu, which may be changed from GAA to GAG.

[0024] In the case of the codons for Met (AUG) and Trp (UGG) there is however no possibility of modifying the sequence.

[0025] The substitutions listed above may be used individually and in all possible combinations in order to increase the G/C content of a modified mRNA compared to the original sequence. Thus, for example all codons for Thr occurring in the original (wild type) sequence can be altered to ACC (or ACG). Preferably, however, combinations of the substitution possibilities given above are employed, for example: substitution of all codons coding in the original sequence for Thr to ACC (or ACG) and substitution of all codons coding for Ser to UCC (or UCG or AGC); substitution of all codons coding in the original sequence for Ile to AUC and substitution of all codons coding for Lys to AAG and substitution of all codons coding originally for Tyr to UAC; substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding for Glu to GAG and substitution of all codons coding for Ala to GCC (or GCG) and substitution of all codons coding for Arg to CGC (or CGG); substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding for Glu to GAG and substitution of all codons coding for Ala to GCC (or GCG) and substitution of all codons coding for Gly to GGC (or GGG) and substitution of all codons coding for Asn to AAC; substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding for Phe to UUC and substitution of all codons coding for Cys to UGC and substitution of all codons coding for Leu to CUG (or CUC) and substitution of all codons coding for Gln to CAG and substitution of all codons encoding Pro to CCC (or CCG); etc.

[0026] Preferably the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased by at least 7% , more preferably by at least 15%, and particularly preferably by at least 20% compared to the G/C content of the coded region of the wild type mRNA encoding for the polypeptide.

[0027] In this connection it is particularly preferred to maximise the G/C content of the modified mRNA as compared to that of the wild type sequence. For some applications, it may be particularly advantageous to maximise the G/C content of the modified mRNA in the region encoding the at least one peptide or polypeptide.

[0028] In accordance with the invention, a further modification of the mRNA comprised in the pharmaceutical composition of the present invention is based on an understanding that the translational efficiency is also affected by the relative abundance of different tRNAs in various cells. A high frequency of so-called "rare" codons in an RNA sequence, which are recognized by relatively rare tRNAs, tends to decrease the translational efficiency of the corresponding mRNA, whereas a high frequency of codons recognized by relatively abundant tRNAs tends to enhance the translational efficiency of a corresponding mRNA.

[0029] Thus, according to the invention, the modified mRNA (which is contained in the pharmaceutical composition) comprises a region coding for the peptide or polypeptide which is changed compared to the corresponding region of the wild type mRNA so as to replace at least one codon of the wild type sequence that is recognized by a rare cellular tRNA with a codon recognized by an abundant cellular tRNA, wherein the abundant and rare cellular tRNAs recognize the same amino acid. In other words, the substituted codon in the modified mRNA, which is recognized by a relatively frequent tRNA, encodes the same amino acid as the wild type (unmodified) codon.

[0030] Through such modifications, the RNA sequences are modified so that codons are inserted/substituted that are recognized by abundantly expressed cellular tRNAs. Modifications directed to altering codon usage in a nucleic acid sequence to optimise expression levels of polypeptides encoded therefrom are generally referred to in the art as "codon optimisation".

[0031] Those tRNAs which are abundant or rare in a particular cell are known to a person skilled in the art; see for example Akashi, Curr. Opin. Genet. Dev. 2001, 11(6): 660-666. Each organism has a preferred choice of nucleotide or codon usage to encode any particular amino acid. Different species vary in their codon preferences for

translating mRNA into protein. The codon preferences of a particular species in which a modified mRNA of the present invention is to be expressed will, therefore, at least in part dictate the parameters of codon optimisation for a nucleic acid sequence.

[0032] By means of this modification, according to the invention all codons of the wild type sequence that are recognized by a relatively rare tRNA in a cell may in each case be replaced by a codon that is recognized by a relatively abundant tRNA. As described herein, however, the coding sequence of the peptide or polypeptide is preserved. That is, a relatively abundant tRNA species, which replaces a relatively rare tRNA species in a modified mRNA of the invention, recognizes an amino acid identical to that recognized by the rare tRNA species.

[0033] According to the invention, it is particularly preferred to couple the sequential increase in the G/C fraction of a modified mRNA (particularly, for example, a maximally modified G/C content), with an increase in the number of codons recognized by abundant tRNAs, wherein the amino acid sequence of the peptide or polypeptide (one or more) encoded by the mRNA remains unaltered. This preferred embodiment provides a particularly preferred mRNA species, possessing properties of efficient translation and improved stability. Such preferred mRNA species are well suited, for example, for the pharmaceutical compositions of the present invention.

[0034] Sequences of eukaryotic mRNAs frequently include destabilising sequence elements (DSE) to which signal proteins can bind and thereby regulate the enzymatic degradation of the mRNA *in vivo*. Accordingly, for the further stabilisation of a modified mRNA of the invention, which may be a component of a pharmaceutical composition of the invention, one or more changes may be made in the wild type mRNA sequence encoding the at least one peptide or polypeptide, so as to reduce the number of destabilising sequence elements present. In accordance with the invention, DSEs located anywhere in an mRNA, including the coding region and in the non-translated regions (3' and/or 5' UTR), may be mutated or changed to generate a modified mRNA having improved properties.

[0035] Such destabilising sequences are for example AU-rich sequences ("AURES") that occur in 3'-UTR regions of a number of unstable mRNAs (Caput et al., Proc. Natl. Acad.

Sci. USA 1986, 83: 1670-1674). The RNA molecules contained in the pharmaceutical composition according to the invention are therefore preferably altered as compared to the wild type mRNA so as to reduce the number of or eliminate these destabilising sequences. Such an approach also applies to those sequence motifs recognised by potential endonucleases. Such sequences include, for example, GAACAAG, which is found in the 3'UTR of the gene encoding the transferring receptor (Binder et al., EMBO J. 1994, 13: 1969-1980). Sequence motifs recognized by endonucleases are also preferably reduced in number or eliminated in the modified mRNA of the pharmaceutical composition according to the invention.

[0036] Various methods are known to the person skilled in the art that are suitable for the substitution of codons in the modified mRNA according to the invention. In the case of relatively short coding regions (that code for biologically active or antigenic peptides), the whole mRNA may, for example, be chemically synthesised using standard techniques.

[0037] Preferably, however, base substitutions are introduced using a DNA matrix for the production of modified mRNA with the aid of techniques routinely employed in targeted mutagenesis; see Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3rd Edition, Cold Spring Harbor, NY, 2001.

[0038] In this method, a corresponding DNA molecule is therefore transcribed in vitro for the production of the mRNA. This DNA matrix has a suitable promoter, for example a T7 or SP6 promoter, for *in vitro* transcription, followed by the desired nucleotide sequence for the mRNA to be produced and a termination signal for the *in vitro* transcription. According to the invention the DNA molecule that forms the matrix of the RNA construct to be produced is prepared as part of a plasmid replicable in bacteria, wherein the plasmid is replicated or amplified during the course of bacterial replication and subsequently isolated by standard techniques. Plasmids suitable for use in the present invention include, but are not limited to pT7Ts (GenBank Accession No. U26404; Lai et al., Development 1995, 121: 2349-2360), the pGEM® series, for example pGEM®-1 (GenBank Accession No. X65300; from Promega) and pSP64 (GenBank-Accession No. X65327); see also Mezei and Storts, Purification of PCR Products, in:

Griffin and Griffin (Eds.), PCR Technology: Current Innovation, CRC Press, Boca Raton, FL, 2001.

[0039] Thus, by using short synthetic DNA oligonucleotides that comprise short single-strand transitions at the corresponding cleavage sites, or by means of genes produced by chemical synthesis, the desired nucleotide sequence can be cloned into a suitable plasmid by molecular biology methods known to the person skilled in the art (see Maniatis et al., above). The DNA molecule is then excised from the plasmid, in which it may be present as a single copy or multiple copies, by digestion with restriction endonucleases.

[0040] The modified mRNA that is contained in the pharmaceutical composition according to the invention may furthermore have a 5′ cap structure (a modified guanosine nucleotide). Examples of suitable cap structures include, but are not limited to m7G(5')ppp (5'(A,G(5')ppp(5')A and G(5')ppp(5')G.

[0041] According to a further preferred embodiment of the present invention the modified mRNA comprises a poly-A tail of at least 50 nucleotides, preferably at least 70 nucleotides, more preferably at least 100 nucleotides and particularly preferably at least 200 nucleotides.

[0042] For efficient translation of the mRNA an productive binding of the ribosomes to the ribosome binding site [Kozak sequence: GCCGCCACCAUGG (SEQ ID NO: 13), the AUG forms the start codon] is generally required. In this regard it has been established that an increased A/U content around this site facilitates more efficient ribosome binding to the mRNA.

[0043] In addition, it is possible to introduce one or more so-called IRES ("internal ribosomal entry site") into the modified mRNA. An IRES may act as the sole ribosome binding site, or may serve as one of the ribosome binding sites of an mRNA. An mRNA comprising more than one functional ribosome binding site may encode several peptides or polypeptides that are translated independently by the ribosomes ("multicistronic mRNA"). Examples of IRES sequences that can be used according to the invention include without limitation, those from picornaviruses (e.g. FMDV), pest viruses (CFFV),

polio viruses (PV), encephalomyocarditis viruses (ECMV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses (HCV), classical swine fever viruses (CSFV), murine leukemia virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV).

[0044] According to a further preferred embodiment of the present invention the modified mRNA comprises in the 5′ non-translated and/or 3′ non-translated regions stabilisation sequences that are capable of increasing the half-life of the mRNA in the cytosol.

[0045] These stabilisation sequences may exhibit 100% sequence homology with naturally occurring sequences that are present in viruses, bacteria and eukaryotic cells, or may be derived from such naturally occurring sequences (i.e., may comprise, e.g., mutations substitutions, or deletions in these sequences). Stabilising sequences that may be used in the present invention include, by way of non-limiting example, the untranslated sequences (UTR) of the β-globin gene of *Homo sapiens* or *Xenopus laevis*. Another example of a stabilisation sequence has the general formula (C/U)CCAN$_x$CCC(U/A)Py$_x$UC(C/U)CC, which is contained in the 3'UTR of the very stable mRNAs that encode α-globin, α-(I)-collagen, 15-lipoxygenase, or tyrosine hydroxylase (C.F. Holcik et al., Proc. Natl. Acad. Sci. USA 1997, 94: 2410-2414). Obviously such stabilisation sequences may be used individually or in combination, as well as in combination with other stabilisation sequences known to a person skilled in the art.

[0046] For the further stabilisation of the modified mRNA it is preferred that the modified mRNA comprises at least one analogue of a naturally occurring nucleotide. This approach is based on the understanding that RNA-decomposing enzymes present in a cell preferentially recognise RNA comprising naturally occurring nucleotides as a substrate. The insertion of nucleotide analogues into an RNA molecule, therefore, retards decomposition of the RNA molecule so modified, whereas the effect of such analogs on translational efficiency, particularly when inserted into the coding region of the mRNA, may result in either an increase or decrease in translation of the modified RNA molecule.

[0047] The following is a non-limiting list of nucleotide analogues that can be used in accordance with the invention: phosphorus amidates, phosphorus thioates, peptide nucleotides, methylphosphonates, 7-deazaguanosine, 5-methylcytosine and inosine. The preparation of such analogues is known to the person skilled in the art, for example from US Patent Nos. 4,373,071, US 4,401,796, US 4,415,732, US 4,458,066, US 4,500,707, US 4,668,777, US 4,973,679, US 5,047,524, US 5,132,418, US 5,153,319, US 5,262,530 and 5,700,642. According to the invention such analogues may be present in non-translated and/or translated regions of the modified mRNA.

[0048] Furthermore the effective transfer of the modified mRNA into the cells to be treated or into the organism to be treated may be improved if the modified mRNA is associated with a cationic peptide or protein, or is bound thereto. In particular in this connection the use of protamine as polycationic, nucleic acid-binding protein is particularly effective. It is also possible to use other cationic peptides or proteins such as poly-L-lysine or histones. Procedures for stabilising mRNA are described in EP-A-1083232, whose relevant disclosure is incorporated herein in its entirety.

[0049] For gene therapy applications, for example, wherein a pharmaceutical composition of the invention is used, the modified mRNA therein codes for at least one biologically active peptide or polypeptide that is not formed or is only insufficiently or defectively formed in the patient to be treated. Administration of a modified mRNA encoding the at least one biologically active peptide or polypeptide or a composition thereof to such a patient, therefore, at least partially restores the expression and/or activity of the at least one biologically active peptide or polypeptide in the patient and thereby complements the patient's genetic defect. The direct introduction of a normal, functional gene into a living animal has been studied as a means for replacing defective genetic information. In such studies, nucleic acid sequences are introduced directly into cells of a living animal. The following references pertain to methods for the direct introduction of nucleic acid sequences into a living animal: Nabel et al., (1990) Science 249:1285-1288; Wolfe et al., (1990) Science 247:1465-1468; Acsadi et al. (1991) Nature 352:815-818; Wolfe et al. (1991) BioTechniques 11(4):474-485; and Felgner and Rhodes, (1991) Nature 349:351-352, which are incorporated herein by reference.

[0050] Accordingly, examples of polypeptides coded by a modified mRNA of the invention include, without limitation, dystrophin, the chloride channel, which is defectively altered in cystic fibrosis; enzymes that are lacking or defective in metabolic disorders such as phenylketonuria, galactosaemia, homocystinuria, adenosine deaminase deficiency, etc.; enzymes that are involved in the synthesis of neurotransmitters such as dopamine, norepinephrine and GABA, in particular tyrosine hydroxylase and DOPA decarboxylase, and α-1-antitrypsin, etc. Pharmaceutical compositions of the invention may also be used to effect expression of cell surface receptors and/or binding partners of cell surface receptors if the modified mRNA contained therein encodes for such biologically active proteins or peptides. Examples of such proteins that act in an extracellular manner or that bind to cell surface receptors include for example tissue plasminogen activator (TPA), growth hormones, insulin, interferons, granulocyte-macrophage colony stimulating factor (GM-CFS), and erythropoietin (EPO), etc. By choosing suitable growth factors, the pharmaceutical composition of the present invention may, for example, be used for tissue regeneration. In this way diseases that are characterised by tissue degeneration, for example neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, etc. and other degenerative conditions, such as arthrosis, can be treated. In these cases the modified mRNA, in particular that contained in the pharmaceutical composition of the present invention, preferably encodes, without limitation, a TGF-β family member, EGF, FGF, PDGF, BMP, GDNF, BDNF, GDF and neurotrophic factors such as NGF, neutrophines, etc.

[0051] A further area of application of the present invention is vaccination, *i.e.* the use of a modified mRNA for inoculation or the use of a pharmaceutical composition comprising a modified mRNA as an inoculating agent, or the use of a modified mRNA in the preparation of the pharmaceutical composition for inoculation purposes. Vaccination is based on introducing an antigen into an organism or subject, in particular into a cell of the organism or subject. In the context of the present invention, the genetic information encoding the antigen is introduced into the organism or subject in the form of a modified mRNA encoding the antigen. The modified mRNA contained in the pharmaceutical composition is translated into the antigen, *i.e.* the polypeptide or antigenic peptide coded by the modified mRNA is expressed, and an immune response directed against the polypeptide or antigenic peptide is stimulated. For vaccination against a pathogenic organism, e.g., a virus, a bacterium, or a protozoan, a surface

antigen of such an organism maybe used as an antigen against which an immune response is elicited. In the context of the present invention, a pharmaceutical composition comprising a modified mRNA encoding such a surface antigen may be used as a vaccine. In applications wherein a genetic vaccine is used for treating cancer, the immune response is directed against tumour antigens by generating a modified mRNA encoding a tumour antigen(s), in particular a protein which is expressed exclusively on cancer cells. Such a modified mRNA encoding a tumour antigen may be used alone or as a component of a pharmaceutical composition according to the invention, wherein administration of either the modified mRNA or a composition thereof results in expression of the cancer antigen(s) in the organism. An immune response to such a vaccine would, therefore, confer to the vaccinated subject a degree of protective immunity against cancers associated with the immunizing cancer antigen. Alternatively, such measures could be used to vaccinate a cancer patient with a modified mRNA encoding a tumour antigen(s) expressed on the patient's cancer cells so as to stimulate the cancer patient's immune response to attack any cancer cells expressing the encoded antigen.

[0052] In its use as a vaccine the pharmaceutical composition according to the invention is suitable in particular for the treatment of cancers (in which the modified mRNA codes for a tumour-specific surface antigen (TSSA), for example for treating malignant melanoma, colon carcinoma, lymphomas, sarcomas, small-cell lung carcinomas, blastomas, etc. A non-limiting list of specific examples of tumour antigens include, *inter alia*, 707-AP, AFP, ART-4, BAGE, β-catenin/m, Bcr-abl, CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gp100, HAGE, HER-2/neu, HLA-A*0201-R170I, HPV-E7, HSP70-2M, HAST-2, hTERT (or hTRT), iCE, KIAA0205, LAGE, LDLR/FUT, MAGE, MART-1/melan-A, MC1R, myosin/m, MUC1, MUM-1, -2, -3, NA88-A, NY-ESO-1, p190 minor bcr-abl, Pml/RARα, PRAME, PSA, PSM, RAGE, RU1 or RU2, SAGE, SART-1 or SART-3, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2 and WT1. In addition to the above application, the pharmaceutical composition of the invention may be used to treat infectious diseases, for example, viral infectious diseases such as AIDS (HIV), hepatitis A, B or C, herpes, herpes zoster (chicken pox), German measles (rubella virus), yellow fever, dengue fever etc. (flavi viruses), flu (influenza viruses), haemorrhagic infectious diseases (Marburg or Ebola viruses), bacterial infectious diseases such as

Legionnaires' disease (Legionella), gastric ulcer (Helicobacter), cholera (Vibrio), *E. coli* infections, staphylococcal infections, salmonella infections or streptococcal infections, tetanus (Clostridium tetani), or protozoan infectious diseases (malaria, sleeping sickness, leishmaniasis, toxoplasmosis, i.e. infections caused by plasmodium, trypanosomes, leishmania and toxoplasma). Preferably also in the case of infectious diseases the corresponding surface antigens with the strongest antigenic potential are encoded by the modified mRNA. With the aforementioned genes of pathogenic vectors or organisms, in particular in the case of viral genes, this is typically a secreted form of a surface antigen. Moreover, according to the invention mRNAs preferably coding for polypeptides are employed, because polypeptides generally comprise multiple epitopes (polyepitopes). Polypeptides comprising polyepitopes include but are not limited to, surface antigens of pathogenic vectors or organisms, or of tumour cells, preferably secreted protein forms.

[0053] Moreover, the modified mRNA according to the invention may comprise in addition to the antigenic or therapeutically active peptide or polypeptide, at least one further functional region that encodes, for example, a cytokine that promotes the immune response (e.g., a monokine, lymphokine, interleukin or chemokine, such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, INF-α, INF-γ, GM-CFS, LT-α or growth factors such as hGH).

[0054] Furthermore, in order to increase immunogenicity, the pharmaceutical composition according to the invention may contain one or more adjuvants. The term "adjuvant" is understood in this context to denote any chemical or biological compound that promotes or augments a specific immune response. Various mechanisms may be involved in this connection, depending on the various types of adjuvants. For example, compounds that promote endocytosis of the modified mRNA contained in the pharmaceutical composition by dentritic cells (DC) form a first class of usable adjuvants. Other compounds that activate or accelerate maturation of DC (for example, lipopolysaccharides, TNF-α or CD40 ligand) comprise a second class of suitable adjuvants. In general, any agent which is recognized as a potential "danger signal" by the immune system (LPS, GP96, oligonucleotides with the CpG motif) or cytokines such as GM-CSF, may be used as an adjuvant. Co-administration of an adjuvant enhances an immune response generated against an antigen encoded by the modified mRNA. The aforementioned cytokines are particularly preferred in this aspect. Other known

adjuvants include aluminium hydroxide, and Freund's adjuvant, as well as the aforementioned stabilising cationic peptides or polypeptides such as protamine. In addition, lipopeptides such as Pam3Cys are also particularly suitable for use as adjuvants in the pharmaceutical composition of the present invention; see Deres et al, Nature 1989, 342: 561-564.

[0055] The pharmaceutical composition according to the invention comprises, in addition to the modified mRNA, a pharmaceutically compatible carrier and/or a pharmaceutically compatible vehicle. Appropriate methods for achieving a suitable formulation and preparation of the pharmaceutical composition according to the invention are described in "Remington's Pharmaceutical Sciences" (Mack Pub. Co., Easton, PA, 1980), which is herein incorporated by reference in its entirety. For parenteral administration suitable carriers include for example sterile water, sterile saline solutions, polyalkylene glycols, hydrogenated naphthalene and in particular biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxypropylene copolymers. Compositions according to the invention may contain fillers or substances such as lactose, mannitol, substances for the covalent coupling of polymers such as for example polyethylene glycol to inhibitors according to the invention, complexing with metal ions or incorporation of materials in or on special preparations of polymer compound, such as for example polylactate, polyglycolic acid, hydrogel or on liposomes, microemulsions, microcells, unilamellar or multilamellar vesicles, erythrocyte fragments or spheroplasts. The respective modifications of the compositions are chosen depending on physical properties such as, for example, solubility, stability, bioavailability or degradability. Controlled or constant release of the active component according to the invention in the composition includes formulations based on lipophilic depot substances (for example fatty acids, waxes or oils). Coatings of substances or compositions according to the invention containing such substances, namely coatings with polymers (for example poloxamers or poloxamines), are also disclosed within the scope of the present invention. Moreover substances or compositions according to the invention may contain protective coatings, for example protease inhibitors or permeability enhancers. Preferred carriers are typically aqueous carrier materials, in which water for injection (WFI) or water buffered with phosphate, citrate or acetate, etc., is used, and the pH is typically adjusted to 5.0 to 8.0, preferably 6.0 to 7.0. The carrier or the vehicle will in addition preferably contain salt constituents,

for example sodium chloride, potassium chloride or other components that for example make the solution isotonic. In addition the carrier or the vehicle may contain, besides the aforementioned constituents, additional components such as human serum albumin (HSA), polysorbate 80, sugars or amino acids.

[0056] The concentration of the modified mRNA in such formulations may therefore vary within a wide range from 1 μg to 100 mg/ml. The pharmaceutical composition according to the invention is preferably administered parenterally, for example intravenously, intraarterially, subcutaneously or intramuscularly to the patient. It is also possible to administer the pharmaceutical composition topically or orally.

[0057] The invention thus also provides a method for the treatment of the aforementioned medical conditions or an inoculation method for the prevention of the aforementioned conditions, which comprises the administration of the pharmaceutical composition according to the invention to a subject or patient, in particular a human patient.

[0058] A typical regimen for preventing, suppressing, or treating a pathology related to a viral, bacterial, or protozoan infection, may comprise administration of an effective amount of a vaccine composition as described herein, administered as a single treatment, or repeated as enhancing or booster dosages, over a period up to and including between one week and about 24 months, or any range or value therein.

[0059] According to the present invention, an "effective amount" of a vaccine composition is one that is sufficient to achieve a desired biological effect. It is understood that nature and manner of the administration and the effective dosage may be determined by a medical practitioner based on a number of variables including the age, sex, health, and weight of the recipient, the medical condition to be treated and its stage of progression, the kind of concurrent treatment, if any, frequency of treatment, and the nature of the desired outcome. The ranges of effective doses provided below are not intended to limit the invention, but are provided as representative preferred dose ranges. However, the most preferred dosage will be tailored to the individual subject, as is understood and determinable by one of skill in the art, without undue experimentation. See, e.g., Berkow et al., eds., The Merck Manual, 16th edition, Merck and Co., Rahway,

N.J., 1992; Goodman et al., eds., Goodman and Gilman's The Pharmacological Basis of Therapeutics, 8th edition, Pergamon Press, Inc., Elmsford, N.Y., (1990); Avery's Drug Treatment: Principles and Practice of Clinical Pharmacology and Therapeutics, 3rd edition, ADIS Press, LTD., Williams and Wilkins, Baltimore, Md. (1987), Ebadi, Pharmacology, Little, Brown and Co., Boston, Mass. (1985); and Katzung, ed. Basic and Clinical Pharmacology, Fifth Edition, Appleton and Lange, Norwalk, Conn. (1992), which references and references cited therein, are entirely incorporated herein by reference.

[0060] The present invention relates to the use of genetic material (e.g., nucleic acid sequences) as immunizing agents. In one aspect, the present invention relates to the introduction of exogenous or foreign modified DNA or RNA molecules into an individual's tissues or cells, wherein these molecules encode an exogenous protein capable of eliciting an immune response to the protein. The exogenous nucleic acid sequences may be introduced alone or in the context of an expression vector wherein the sequences are operably linked to promoters and/or enhancers capable of regulating the expression of the encoded proteins. The introduction of exogenous nucleic acid sequences may be performed in the presence of a cell stimulating agent capable of enhancing the uptake or incorporation of the nucleic acid sequences into a cell. Such exogenous nucleic acid sequences may be administered in a composition comprising a biologically compatible or pharmaceutically acceptable carrier. The exogenous nucleic acid sequences may be administered by a variety of means, as described herein, and well known in the art.

[0061] Such methods may be used to elicit immunity to a pathogen, absent the risk of infecting an individual with the pathogen. The present invention may be practiced using procedures known in the art, such as those described in PCT International Application Number PCT/US90/01515, wherein methods for immunizing an individual against pathogen infection by directly injecting polynucleotides into the individual's cells in a single step procedure are presented.

[0062] In one aspect, the present invention relates to methods for eliciting immune responses in an individual or subject which can protect the individual from pathogen infection. Accordingly, genetic material that encodes an immunogenic protein is

introduced into a subject's cells either *in vivo* or *ex vivo*. The genetic material is expressed by these cells, thereby producing immunogenic target proteins capable of eliciting an immune response. The resulting immune response is broad based and involves activation of the humoral immune response and both arms of the cellular immune response.

[0063] This approach is useful for eliciting a broad range of immune responses against a target protein. Target proteins may be proteins specifically associated with pathogens or the individual's own "abnormal" or infected cells. Such an approach may be used advantageously to immunize a subject against pathogenic agents and organisms such that an immune response against a pathogen protein provides protective immunity against the pathogen. This approach is particularly useful for protecting an individual against infection by non-encapsulated intracellular pathogens, such as a virus, which produce proteins within the host cells. The immune response generated against such proteins is capable of eliminating infected cells with cytotoxic T cells (CTLs).

[0064] The immune response elicited by a target protein produced by vaccinated cells in a subject is a broad-based immune response which includes B cell and T cell responses, including CTL responses. It has been observed that target antigen produced within the cells of the host are processed intracellularly into small peptides, which are bound by Class I MHC molecules and presented in the context of Class I on the cell surface. The Class I MHC-target antigen complexes are capable of stimulating $CD8^+$ T cells, which are predominantly CTLs. Notably, genetic immunization according to the present invention is capable of eliciting CTL responses (killer cell responses).

[0065] The CTL response is crucial in protection against pathogens such as viruses and other intracellular pathogens which produce proteins within infected cells. Similarly, the CTL response can be utilized for the specific elimination of deleterious cell types, which may express aberrant cell surface proteins recognizable by Class I MHC molecules.

[0066] The genetic vaccines of the present invention may be administered to cells in conjunction with compounds that stimulate cell division and facilitate uptake of genetic constructs. This step provides an improved method of direct uptake of genetic material. Administration of cell stimulating compounds results in a more effective immune

response against the target protein encoded by the genetic construct.

[0067] According to the present invention, modified DNA or mRNA that encodes a target protein is introduced into the cells of an individual where it is expressed, thus producing the target protein. The modified DNA or RNA may be operably linked to regulatory elements (e.g., a promoter) necessary for expression in the cells of the individual. Other elements known to skilled artisans may also be included in genetic constructs of the invention, depending on the application.

[0068] As used herein, the term "genetic construct" refers to the modified DNA or mRNA molecule that comprises a nucleotide sequence which encodes the target protein and which may include initiation and termination signals operably linked to regulatory elements including a promoter and polyadenylation signal (for modified DNA) capable of directing expression in the cells of the vaccinated individual. As used herein, the term "expressible form" refers to gene constructs which contain the necessary regulatory elements operably linked to a coding sequence of a target protein, such that when present in the cell of the individual, the coding sequence is expressed. As used herein, the term "genetic vaccine" refers to a pharmaceutical preparation that comprises a genetic construct.

[0069] The present invention provides genetic vaccines, which include genetic constructs comprising DNA or RNA which encode a target protein. As used herein, the term "target protein" refers to a protein capable of eliciting an immune response. The target protein is an immunogenic protein derived from the pathogen or undesirable cell-type, such as an infected or transformed cell. In accordance with the invention, target proteins may be pathogen-associated proteins or tumour-associated proteins. The immune response directed against the target protein protects the individual against the specific infection or disease with which the target protein is associated. For example, a genetic vaccine comprising a modified DNA or RNA molecule that encodes a pathogen-associated target protein is used to elicit an immune response that will protect the individual from infection by the pathogen.

[0070] DNA and RNA-based vaccines and methods of use are described in detail in several publications, including Leitner et al. (1999, Vaccines 18:765-777),

Nagashunmugam et al. (1997, AIDS 11:1433-1444), and Fleeton et al. (2001, J Infect Dis 183:1395-1398) the entire contents of each of which is incorporated herein by reference.

[0071] In order to test expression, genetic constructs can be tested for expression levels *in vitro* using cells maintained in culture, which are of the same type as those to be vaccinated. For example, if the genetic vaccine is to be administered into human muscle cells, muscle cells grown in culture such as solid muscle tumor cells of rhabdomyosarcoma may be used as an *in vitro* model for measuring expression levels. One of ordinary skill in the art could readily identify a model *in vitro* system which may be used to measure expression levels of an encoded target protein.

[0072] In accordance with the invention, multiple inoculants can be delivered to different cells, cell types, or tissues in an individual. Such inoculants may comprise the same or different nucleic acid sequences of a pathogenic organism. This allows for the introduction of more than a single antigen target and maximizes the chances for developing immunity to the pathogen in a vaccinated subject.

[0073] According to the invention, the genetic vaccine may be introduced *in vivo* into cells of an individual to be immunized or *ex vivo* into cells of the individual which are re-implanted after incorporation of the genetic vaccine. Either route may be used to introduce genetic material into cells of an individual. As described herein above, preferred routes of administration include intramuscular, intraperitoneal, intradermal, and subcutaneous injection. Alternatively, the genetic vaccine may be introduced by various means into cells isolated from an individual. Such means include, for example, transfection, electroporation, and microprojectile bombardment. These methods and other protocols for introducing nucleic acid sequences into cells are known to and routinely practiced by skilled practitioners. After the genetic construct is incorporated into the cells, they are re-implanted into the individual. Prior to re-implantation, the expression levels of a target protein encoded by the genetic vaccine may be assessed. It is contemplated that otherwise non-immunogenic cells that have genetic constructs incorporated therein can be implanted into autologous or heterologous recipients.

[0074] The genetic vaccines according to the present invention comprise about 0.1 to

about 1000 micrograms of nucleic acid sequences (i.e., DNA or RNA). In some preferred embodiments, the vaccines comprise about 1 to about 500 micrograms of nucleic acid sequences. In some preferred embodiments, the vaccines comprise about 25 to about 250 micrograms of nucleic acid sequences. Most preferably, the vaccines comprise about 100 micrograms nucleic acid sequences.

[0075] The genetic vaccines according to the present invention are formulated according to the mode of administration to be used. One having ordinary skill in the art can readily formulate a genetic vaccine that comprises a genetic construct. In cases where intramuscular injection is the chosen mode of administration, for example, an isotonic formulation is generally used. As described in detail herein above, additives for isotonicity can include sodium chloride, dextrose, mannitol, sorbitol and lactose. Isotonic solutions such as phosphate buffered saline are preferred. Stabilizers can include gelatin and albumin.

[0076] In some embodiments of the invention, the individual is administered a series of vaccinations to produce a comprehensive immune response. According to this method, at least two and preferably four injections are given over a period of time. The period of time between injections may include from 24 hours apart to two weeks or longer between injections, preferably one week apart. Alternatively, at least two and up to four separate injections may be administered simultaneously to different parts of the body.

[0077] While this disclosure generally discusses immunization or vaccination in the context of prophylactic methods of protection, the terms "immunizing" or "vaccinating" are meant to refer to both prophylactic and therapeutic methods. Thus, a method for immunizing or vaccinating includes both methods of protecting an individual from pathogen challenge, as well as methods for treating an individual suffering from pathogen infection. Accordingly, the present invention may be used as a vaccine for prophylactic protection or in a therapeutic manner; that is, as a reagent for immunotherapeutic methods and preparations.

[0078] The amount of a modified nucleic acid sequence generated using the methods of the invention which provides a therapeutically effective dose in the treatment of a patient with, for example, cancer or a pathogen-related disorder can be determined by standard

clinical techniques based on the present description. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each subject's circumstances. However, suitable dosage ranges for intravenous administration are generally directed to achieve a concentration of about 20-500 micrograms of polypeptide encoded by the modified nucleic acid per kilogram body weight. Suitable dosage ranges for intranasal administration are generally directed to achieve a concentration of about 0.01 pg to 1 mg of polypeptide encoded by the modified nucleic acid per kg body weight. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

[0079] The compositions comprising the modified nucleic acid molecules of the invention can be administered for prophylactic and/or therapeutic treatments. In therapeutic applications, compositions are administered to a patient already suffering from a hyperproliferative disorder (such as, e.g., cancer) in an amount sufficient to cure or at least partially arrest the symptoms of the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective amount or dose." Amounts effective for this use will depend on the severity of the disease and the weight and general state of the patient.

[0080] Compositions comprising modified nucleic acid molecules of the invention can be administered alone, or in combination, and/or in conjunction with known therapeutic agents/compounds used for the treatment of a patient with a particular disorder. For the treatment of a patient with cancer, for example, a composition comprising at least one modified nucleic acid of the invention which encodes a tumour antigen, may be used in conjunction with one or more known cancer therapeutics, such as those described in the Physicians' Desk Reference, 54th Edition (2000) or in Cancer: Principles & Practice of Oncology, DeVita, Jr., Hellman, and Rosenberg (eds.) 2nd edition, Philadelphia, PA: J.B. Lippincott Co., 1985, wherein standard treatment protocols and dosage formulations are presented.

[0081] In addition a method is also provided for determining how to modify the sequence of an mRNA so as to generate a modified mRNA having altered properties,

which may be used alone or in a pharmaceutical composition of the invention. In this connection, and in accordance with the invention, the modification of an RNA sequence is carried out with two different optimisation objectives: to maximize G/C content, and to maximize the frequency of codons that are recognized by abundantly expressed tRNAs. In the first step of the process a virtual translation of an arbitrary RNA (or DNA) sequence is carried out in order to generate the corresponding amino acid sequence. Starting from the amino acid sequence, a virtual reverse translation is performed that provides, based on degeneracy of the genetic code, all of the possible choices for the corresponding codons. Depending on the required optimisation or modification, corresponding selection lists and optimisation algorithms are used for choosing suitable codons. The algorithms are executed on a computer, normally with the aid of suitable software. In accordance with the present invention, a suitable software program comprises a source code of Appendix I Thus, the optimised mRNA sequence is generated and can be output, for example, with the aid of a suitable display device and compared with the original (wild type) sequence. The same also applies with regard to the frequency of the individual nucleotides. The changes compared to the original nucleotide sequence are preferably emphasised. Furthermore, according to a preferred embodiment, naturally occurring stable sequences are incorporated therein to produce an RNA stabilised by the presence of natural sequence motifs. A secondary structural analysis may also be performed that can analyse, on the basis of structural calculations, stabilising and destabilising properties or regions of the RNA.

[0082] Also encompassed by the present invention are modified nucleic acid sequences generated using the above computer-based method. Exemplary modified nucleic acid sequences of the invention include SEQ ID NOs: 3-7, 10 and 11. The present invention also includes pharmaceutical compositions of modified nucleic acid sequences of the invention, including SEQ ID NOs: 3-7, 10 and 11.

BRIEF DESCRIPTION OF THE DRAWINGS

[0083] Fig. 1 shows wild type sequences and modified sequences for the influenza matrix protein.

[0084] Fig. 1A (SEQ ID NO: 1) shows the wild type gene and Fig. 1B (SEQ ID NO: 2) shows the amino acid sequence derived therefrom (1-letter code). Fig. 1C (SEQ ID NO: 3) shows a gene sequence coding for the influenza matrix protein, whose G/C content is increased as compared to that of the wild type sequence. Fig. 1D (SEQ ID NO: 4) shows the sequence of a gene that codes for a secreted form of the influenza matrix protein (including an N-terminal signal sequence), wherein the G/C content of the sequence is increased relative to that of the wild type sequence. Fig. 1E (SEQ ID NO: 5) shows an mRNA coding for the influenza matrix protein, wherein the mRNA comprises stabilising sequences not present in the corresponding wild type mRNA. Fig. 1F (SEQ ID NO: 6) shows an mRNA coding for the influenza matrix protein that in addition to stabilising sequences also contains an increased G/C content. Fig. 1G (SEQ ID NO: 7) likewise shows a modified mRNA that codes for a secreted form of the influenza matrix protein and comprises, as compared to the wild type, stabilising sequences and an elevated G/C content. In Fig. 1A and Figs. 1C to 1G the start and stop codons are shown in bold type. Nucleotides that are changed relative to the wild type sequence of Fig. 1A are shown in capital letters in 1C to 1G.

[0085] Fig. 2 shows wild type sequences and modified sequences according to the invention that encode for the tumour antigen MAGE1.

[0086] Fig. 2A (SEQ ID NO: 8) shows the sequence of the wild type gene and Fig. 2B (SEQ ID NO: 9) shows the amino acid sequence derived therefrom (3-letter code). Fig. 2C (SEQ ID NO: 10) shows a modified mRNA coding for MAGE1, whose G/C content is increased as compared to the wild type. Fig. 2D (SEQ ID NO: 11) shows the sequence of a modified mRNA encoding MAGE1, in which the codon usage has been optimised as frequently as possible with respect to the tRNA present in the cell and to the coding sequence in question. Start and stop codons are shown in each case in bold type.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0087] The following examples describe the invention in more detail and in no way are to be construed as restricting the scope thereof.

Example 1

[0088] As an exemplary embodiment of the process for determining the sequence of a modified mRNA according to the invention, a computer program was established that modifies the nucleotide sequence of an arbitrary mRNA in such a way as to maximise the G/C content of the nucleic acid, and maximise the presence of codons recognized by abundant tRNAs present in a particular cell(s). The computer program is based on an understanding of the genetic code and exploits the degenerative nature of the genetic code. By this means a modified mRNA having desirable properties is obtained, wherein the amino acid sequence encoded by the modified mRNA is identical to that of the unmodified mRNA sequence. Alternatively, the invention may encompass alterations in either the G/C content or codon usage of an mRNA to produce a modified mRNA.

[0089] The source code in Visual Basic 6.0 (program development environment employed: Microsoft Visual Studio Enterprise 6.0 with Servicepack 3) is given in the Appendix I.

Example 2

[0090] An RNA construct with a sequence of the lac-Z gene from *E. coli* optimised with regard to stabilisation and translational efficiency was produced with the aid of the computer program of Example 1. A G/C content of 69% (compared to the wild type sequence of 51%; C.F. Kalnins et al., EMBO J. 1983, 2(4): 593-597) was achieved in this manner. Through the synthesis of overlapping oligonucleotides that comprise the modified sequence, the optimised sequence was produced according to methods known in the art. The terminal oligonucleotides have the following restriction cleavage sites: at the 5′ end an *Eco*RV cleavage site, and at the 3' end a *Bgl*II cleavage site. The modified lacZ sequence was incorporated into the plasmid pT7Ts (GenBank Accession No. U26404; C.F. Lai et al., see above) by digestion with *Eco*RV/*Bgl*II. pT7Ts contains untranslated region sequences from the β-globin gene of *Xenopus laevis* at the 5' and 3' ends. The plasmid was cleaved with the aforementioned restriction enzymes to facilitate insertion of the modified lacZ sequence having compatible 5' and 3' termini.

[0091] The pT7Ts-lac-Z construct was propagated in bacteria and purified by phenol-chloroform extraction. 2 μg of the construct were transcribed *in vitro* using methods known to a skilled artisan and the modified mRNA was produced.

Example 3

[0092] The gene for the influenza matrix protein (wild type sequence, see Fig. 1A; derived amino acid sequence, see Fig. 1B) was optimised with the aid of the computer program according to the invention of Example 1. The G/C-rich sequence variant shown in Fig. 1C (SEQ ID NO: 3) was thereby formed. A G/C-rich sequence coding for a secreted form of the influenza matrix protein, which includes an N-terminal signal sequence was also determined (see Fig. 1D; SEQ ID NO: 4). The secreted form of the influenza matrix protein has the advantage of increased immunogenicity as compared to that of the non-secreted form.

[0093] Corresponding mRNA molecules were designed starting from the optimised sequences. The mRNA for the influenza matrix protein, optimised with regard to G/C content and codon usage, was additionally provided with stabilising sequences in the 5′ region and 3′ region (the stabilisation sequences derive from the 5′-UTRs and 3′-UTRs of the β-globin-mRNA of *Xenopus laevis*; pT7Ts-Vektor in C.F.Lai et al., see above). See also Figs. 1E; SEQ ID NO: 5, which includes only stabilising sequences and 1F; SEQ ID NO: 6, which includes both increased G/C content and stabilising sequences. The mRNA coding for the secreted form of the influenza matrix protein was likewise also sequence optimised in the translated region and provided with the aforementioned stabilising sequences (see Fig. 1G; SEQ ID NO: 7).

Example 4

[0094] The mRNA encoding the tumour antigen MAGE1 was modified with the aid of the computer program of Example 1. The sequence shown in Fig. 2C (SEQ ID NO: 10) was generated in this way, and has a 24% higher G/C content (351 G, 291 C) as compared to the wild type sequence (275 G, 244 G). In addition, by means of alternative codon usage, the wild type sequence was improved with regard to translational efficiency by substituting codons corresponding to tRNAs that are more abundant in a target cell

(see Fig. 2D; SEQ ID NO: 11). The G/C content was likewise raised by 24% by the alternative codon usage.

Appendix I: Source text of a computer program according to the invention

Curevac_Genetic_Controls with the following modules

Name: Curevac_Genetic_Controls.vbp
Code:

```
Type=Control
Reference=*\G{00020430-0000-0000-C000-
000000000046}#2.0#0#..\..\WINNT\System32\STDOLE2.TLB#OLE Automation
Object={0D452EE1-E08F-101A-852E-02608C4D0BB4}#2.0#0; FM20.DLL
Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}#1.2#0; COMDLG32.OCX
UserControl=Curevac_Amino.ctl
Startup="(None)"
HelpFile=""
Title="Curevac Genetic Controls"
Command32=""
Name="Curevac_Genetic_Controls"
HelpContextID="0"
CompatibleMode="1"
MajorVer=1
MinorVer=0
RevisionVer=0
AutoIncrementVer=0
ServerSupportFiles=0
VersionComments="the RNA people"
VersionCompanyName="CureVac GmbH"
VersionFileDescription="Controls for Handling of Nucleotids"
VersionLegalCopyright="by Christian Klump"
VersionLegalTrademarks="Curevac Genetic Controls(tm)"
VersionProductName="Curevac Genetic Controls"
CompilationType=0
OptimizationType=0
FavorPentiumPro(tm)=0
CodeViewDebugInfo=0
NoAliasing=0
BoundsCheck=0
OverflowCheck=0
FlPointCheck=0
FDIVCheck=0
UnroundedFP=0
StartMode=1
Unattended=0
Retained=0
ThreadPerObject=0
MaxNumberOfThreads=1
ThreadingModel=1
DebugStartupOption=1
DebugStartupComponent=Curevac_Amino
```

Name: Curevac_Amino.ctl
Code:

```
VERSION 5.00
Object = "{0D452EE1-E08F-101A-852E-02608C4D0BB4}#2.0#0"; "FM20.DLL"
Begin VB.UserControl Curevac_Amino
  CanGetFocus    =  0  'False
  ClientHeight   =  690
  ClientLeft     =  0
  ClientTop      =  0
  ClientWidth    =  1200
  ScaleHeight    =  690
  ScaleWidth     =  1200
  Begin VB.Line linLower
    X1           =  0
    X2           =  1080
    Y1           =  600
    Y2           =  600
  End
  Begin VB.Line linUpper
    X1           =  120
    X2           =  1080
    Y1           =  0
    Y2           =  0
  End
  Begin MSForms.Label lblAminoAcid
    DragMode     =  1  'Automatic
    Height       =  255
    Left         =  120
    TabIndex     =  1
    Top          =  360
    Width        =  975
    Size         =  "1720;450"
    SpecialEffect  =  1
    FontName     =  "Lucida Sans Unicode"
    FontEffects    =  1073741826
    FontHeight     =  165
    FontCharSet    =  0
    FontPitchAndFamily=  2
    ParagraphAlign  =  3
  End
  Begin MSForms.Label lblTriplett
    DragMode     =  1  'Automatic
    Height       =  255
    Left         =  120
    TabIndex     =  0
    Top          =  120
    Width        =  975
    Size         =  "1720;450"
```

```
    SpecialEffect   =   1
    FontHeight      =   165
    FontCharSet     =   0
    FontPitchAndFamily=   2
    ParagraphAlign  =   3
  End
End
Attribute VB_Name = "Curevac_Amino"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = False
Attribute VB_Exposed = True
Option Explicit

Private msAAShortcut As String
Private msAAName As String
Private msBestTriplett As String
Private msSecondBest As String
Private msThirdBest As String
Private msTriplett As String
Private msBackColor As Long
Private mbShowOriginal As Boolean


Public Enum enuAminoAcid
   amaGlycin
   amaAlanin
   amaValin
   amaLeucin
   amaIsoLeucin
   amaPhenylalanin
   amaTyrosin
   amaTryptophan
   amaAsparaginAcid
   amaAsparagin
   amaGlutaminAcid
   amaGlutamin
   amaSerin
   amaThreonin
   amaCystein
   amaMethionin
   amaProlin
   amaHistidin
   amaLysin
   amaStop
   amaStart
End Enum

Private Sub Main()

   Call UserControl_Resize
```

```
End Sub

Public Sub Translation(ByVal sTriplett As String)

   msTriplett = sTriplett

   Select Case sTriplett
      Case "GGU", "GGC", "GGA", "GGG"
         msAAShortcut = "GLY"
      Case "GCU", "GCC", "GCA", "GCG"
         msAAShortcut = "ALA"
      Case "GUU", "GUC", "GUA", "GUG"
         msAAShortcut = "VAL"
      Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
         msAAShortcut = "LEU"
      Case "AUA", "AUU", "AUC"
         msAAShortcut = "ILE"
      Case "UUU", "UUC"
         msAAShortcut = "PHE"
      Case "UAU", "UAC"
         msAAShortcut = "TYR"
      Case "UGG"
         msAAShortcut = "TRP"
      Case "GAU", "GAC"
         msAAShortcut = "ASP"
      Case "AAU", "AAC"
         msAAShortcut = "ASN"
      Case "GAA", "GAG"
         msAAShortcut = "GLU"
      Case "CAA", "CAG"
         msAAShortcut = "GLN"
      Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
         msAAShortcut = "SER"
      Case "ACA", "ACU", "ACG", "ACC"
         msAAShortcut = "THR"
      Case "UGU", "UGC"
         msAAShortcut = "CYS"
      Case "AUG"
         msAAShortcut = "MET"
      Case "CCA", "CCU", "CCG", "CCC"
         msAAShortcut = "PRO"
      Case "CAU", "CAC"
         msAAShortcut = "HIS"
      Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
         msAAShortcut = "ARG"
      Case "AAA", "AAG"
         msAAShortcut = "LYS"
      Case "UAA", "UAG", "UGA"
         msAAShortcut = "STP"
   End Select
```

```
    Call Synthesis(msAAShortcut)
    Call Show

End Sub

Public Sub Synthesis(ByVal sShortCut As String)

    msAAShortcut = sShortCut 'nur falls die Sub extern angesprochen wird

    Select Case msAAShortcut
      Case "GLY"
        msAAName = "Glycin"
        msBestTriplett = "GGC"
        msSecondBest = "GGG"
        msBackColor = 8421631
      Case "ALA"
        msAAName = "Alanin"
        msBestTriplett = "GCG"
        msSecondBest = "GCC"
        msBackColor = 8454143
      Case "VAL"
        msAAName = "Valin"
        msBestTriplett = "GUC"
        msSecondBest = "GUG"
        msBackColor = 8454016
      Case "LEU"
        msAAName = "Leucin"
        msBestTriplett = "CUG"
        msSecondBest = "CUC"
        msBackColor = 16777088
      Case "ILE"
        msAAName = "Isoleucin"
        msBestTriplett = "AUC"
        msBackColor = 12615935
      Case "PHE"
        msAAName = "Phenylalanin"
        msBestTriplett = "UUC"
        msBackColor = 255
      Case "TYR"
        msAAName = "Tyrosin"
        msBestTriplett = "UAC"
        msBackColor = 4210816
      Case "TRP"
        msAAName = "Tryptophan"
        msBestTriplett = "UGG"
        msBackColor = 4227327
      Case "ASP"
        msAAName = "Asparaginsäure"
        msBestTriplett = "GAC"
        msBackColor = 8388863
      Case "ASN"
```

```
    msAAName = "Asparagin"
    msBestTriplett = "AAC"
    msBackColor = 4227200
  Case "GLU"
    msAAName = "Glutaminsäure"
    msBestTriplett = "GAG"
    msBackColor = 32768
  Case "GLN"
    msAAName = "Glutamin"
    msBestTriplett = "GGC"
    msBackColor = 8421440
  Case "SER"
    msAAName = "Serin"
    msBestTriplett = "AGC"
    msSecondBest = "UCG"
    msThirdBest = "UCC"
    msBackColor = 16512
  Case "THR"
    msAAName = "Threonin"
    msBestTriplett = "ACG"
    msBackColor = 16711680
  Case "CYS"
    msAAName = "Cystein"
    msBestTriplett = "UGC"
    msBackColor = 6932960
  Case "MET"
    msAAName = "Methionin"
    msBestTriplett = "AUG"
    msBackColor = 10417643
  Case "PRO"
    msAAName = "Prolin"
    msBestTriplett = "CCG"
    msSecondBest = "CCC"
    msBackColor = 12898746
  Case "HIS"
    msAAName = "Histidin"
    msBestTriplett = "CAC"
    msBackColor = 12898746
  Case "ARG"
    msAAName = "Arginin"
    msBestTriplett = "CGC"
    msSecondBest = "CGG"
    msBackColor = 6174925
  Case "LYS"
    msAAName = "Lysin"
    msBestTriplett = "AAG"
    msBackColor = 14141641
  Case "STP"
     msAAName = "Stop"
    msBestTriplett = "UGA"
    msSecondBest = "UAG"
```

```
        msBackColor = 11332093
    End Select

End Sub

Public Sub Show()

    lblAminoAcid.Caption = msAAShortcut
    If mbShowOriginal = True Then
        lblTriplett.Caption = msTriplett
    Else
        lblTriplett.Caption = msBestTriplett
    End If
    lblAminoAcid.BackColor = msBackColor
    lblTriplett.BackColor = msBackColor

End Sub

Public Function ChooseBestGC(sShortCut As String)

    Select Case sShortCut
        Case "GLY"
            msAAName = "Glycin"
            msBestTriplett = "GGC"
            msSecondBest = "GGG"
            msBackColor = 8421631
        Case "ALA"
            msAAName = "Alanin"
            msBestTriplett = "GCG"
            msSecondBest = "GCC"
            msBackColor = 8454143
        Case "VAL"
            msAAName = "Valin"
            msBestTriplett = "GUC"
            msSecondBest = "GUG"
            msBackColor = 8454016
        Case "LEU"
            msAAName = "Leucin"
            msBestTriplett = "CUG"
            msSecondBest = "CUC"
            msBackColor = 16777088
        Case "ILE"
            msAAName = "Isoleucin"
            msBestTriplett = "AUC"
            msBackColor = 12615935
        Case "PHE"
            msAAName = "Phenylalanin"
            msBestTriplett = "UUC"
            msBackColor = 255
        Case "TYR"
            msAAName = "Tyrosin"
```

```
    msBestTriplett = "UAC"
    msBackColor = 4210816
Case "TRP"
    msAAName = "Tryptophan"
    msBestTriplett = "UGG"
    msBackColor = 4227327
Case "ASP"
    msAAName = "Asparaginsäure"
    msBestTriplett = "GAC"
    msBackColor = 8388863
Case "ASN"
    msAAName = "Asparagin"
    msBestTriplett = "AAC"
    msBackColor = 4227200
Case "GLU"
    msAAName = "Glutaminsäure"
    msBestTriplett = "GAG"
    msBackColor = 16711808
Case "GLN"
    msAAName = "Glutamin"
    msBestTriplett = "GGC"
    msBackColor = 12632256
Case "SER"
    msAAName = "Serin"
    msBestTriplett = "AGC"
    msSecondBest = "UCG"
    msThirdBest = "UCC"
    msBackColor = 8421504
Case "THR"
    msAAName = "Threonin"
    msBestTriplett = "ACG"
    msBackColor = &HD4C0FF
Case "CYS"
    msAAName = "Cystein"
    msBestTriplett = "UGC"
    msBackColor = &HD5C0FF
Case "MET"
    msAAName = "Methionin"
    msBestTriplett = "AUG"
    msBackColor = &HD6C0FF
Case "PRO"
    msAAName = "Prolin"
    msBestTriplett = "CCG"
    msSecondBest = "CCC"
    msBackColor = &HD7C0FF
Case "HIS"
    msAAName = "Histidin"
    msBestTriplett = "CAC"
    msBackColor = &HD8C0FF
Case "ARG"
    msAAName = "Arginin"
```

```
            msBestTriplett = "CGC"
            msSecondBest = "CGG"
            msBackColor = &HD9C0FF
        Case "LYS"
            msAAName = "Lysin"
            msBestTriplett = "AAG"
            msBackColor = &HE0C0FF
        Case "STP"
            msAAName = "Stop"
            msBestTriplett = "UGA"
            msSecondBest = "UAG"
            msBackColor = &HE1C0FF
    End Select


End Function



Public Property Let ShowOriginal(ByVal bNewValue As Boolean)
    mbShowOriginal = bNewValue
End Property

Public Property Get ShowOriginal() As Boolean
    ShowOriginal = mbShowOriginal
End Property

Public Property Get AAShortcut() As String
    AAShortcut = msAAShortcut
End Property

Public Property Get AAName() As String
    AAName = msAAName
End Property

Public Property Get BestTriplett() As String
    BestTriplett = msBestTriplett
End Property

Public Property Get SecondBest() As String
    SecondBest = msSecondBest
End Property

Public Property Get ThirdBest() As String
    ThirdBest = msThirdBest
End Property

Public Property Get BackColor() As String
    BackColor = msBackColor
End Property
```

```
Private Property Get Triplett() As String
   Triplett = msTriplett
End Property

Private Sub UserControl_Resize()

Dim lOuterWidth As Long
Dim lDistance As Long

   lDistance = 30

   lOuterWidth = ScaleWidth + lDistance


   With lblTriplett
      .Top = 0
      .Left = lDistance / 2
      .Width = ScaleWidth - lDistance
      .Height = ScaleHeight / 2 + 30
   End With
   With lblAminoAcid
      .Top = ScaleHeight / 2 - 30
      .Left = lDistance / 2
      .Width = ScaleWidth - lDistance
      .Height = ScaleHeight / 2
   End With
   With linUpper
      .X1 = ScaleLeft
      .Y1 = 0
      .X2 = ScaleWidth
      .Y2 = 0
   End With
   With linLower
      .X1 = ScaleLeft
      .Y1 = ScaleHeight - 20
      .X2 = ScaleWidth
      .Y2 = ScaleHeight - 20
   End With

End Sub
```

Curevac_RNA_Optimizer with following modules

Name: Curevac_RNA_Optimizer.vbp
Code:
Type=Exe

```
Reference=*\G{00020430-0000-0000-C000-
000000000046}#2.0#0#..\..\WINNT\System32\STDOLE2.TLB#OLE Automation
Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}#1.2#0; COMDLG32.OCX
Module=basMain; basMain.bas
Form=mdiMain.frm
Object=*\ACurevac_Genetic_Controls.vbp
Object={38911DA0-E448-11D0-84A3-00DD01104159}#1.1#0; COMCT332.OCX
Object={3B7C8863-D78F-101B-B9B5-04021C009402}#1.2#0; RICHTX32.OCX
Form=frmOutPut.frm
Form=frmStatistics.frm
Module=basPublics; basPublics.bas
Form=frmOptions.frm
Form=frmInput.frm
Form=frmDisplay.frm
Form=frmAbout.frm
Startup="mdiMain"
HelpFile=""
Title="RNA-Optimizer"
Command32=""
Name="RNA_Optimizer"
HelpContextID="0"
CompatibleMode="0"
MajorVer=1
MinorVer=0
RevisionVer=0
AutoIncrementVer=0
ServerSupportFiles=0
VersionComments="the RNA people"
VersionCompanyName="CureVac GmbH"
VersionFileDescription="Application for Optimization of RNA"
VersionLegalCopyright="by Christian Klump"
VersionLegalTrademarks="Curevac RNA-Optimizer(tm)"
VersionProductName="Curevac RNA-Optimizer"
CompilationType=0
OptimizationType=0
FavorPentiumPro(tm)=0
CodeViewDebugInfo=0
NoAliasing=0
BoundsCheck=0
OverflowCheck=0
FlPointCheck=0
FDIVCheck=0
UnroundedFP=0
StartMode=0
Unattended=0
Retained=0
ThreadPerObject=0
MaxNumberOfThreads=1
DebugStartupOption=0
```

Name: basMain.bas
Code:

```
Attribute VB_Name = "basMain"
Option Explicit

Private Type udtAminoAcid
   sShortcut As String
   sName As String
   sBestTriplett As String
   sSecondBest As String
   sThirdBest As String
End Type

Private mastrAminoAcids() As udtAminoAcid

Public Sub RebuildChain(bShowOriginal As Boolean)

Dim iLoop As Integer

   For iLoop = 1 To glAminoCounter
      With frmDisplay.Curevac_Amino1(iLoop)
         .ShowOriginal = bShowOriginal
         .Show
      End With
   Next iLoop

End Sub

Public Sub ResetPublics()
   gsOriginalCode = ""
   gsFormattedCode = ""
   glCodeLength = ""
End Sub


Public Sub BuildAminoChain(ByVal sTempCode As String, bGraphical As Boolean)

Dim lVerticalOffset As Long
Dim lHorizontalOffset As Long
Dim lTop As Long

Dim asTripletts() As String
Dim sAminoChain As String
Dim lLoop As Long

   If bGraphical = True Then
      lTop = frmDisplay.optNucleotids(0).Height + 40
   End If
   ReDim mastrAminoAcids(glAminoCounter)
   asTripletts = Split(sTempCode)
   For lLoop = 0 To glAminoCounter - 1
```

```
      If bGraphical = True Then
         Load frmDisplay.Curevac_Amino1(lLoop + 1)
         With frmDisplay.Curevac_Amino1(lLoop + 1)

            Call .Translation(asTripletts(lLoop))

            .Left = lLoop * .Width - lHorizontalOffset
            .Top = lVerticalOffset + lTop
            If .Left + 2 * .Width > frmDisplay.ScaleWidth Then
               lVerticalOffset = lVerticalOffset + .Height + 150
               lHorizontalOffset = lLoop * .Width + .Width
            End If

            sAminoChain = sAminoChain + .aaShortcut + ","
            .Visible = True
         End With
      Else
         mastrAminoAcids(lLoop) = Translation(asTripletts(lLoop))
'         sTest = sTest + mastrAminoAcids(iLoop).sName + ","
      End If
   Next
   If bGraphical = True Then
      frmDisplay.Show
   Else
      Call frmOutPut.Show
   End If
   Debug.Print sAminoChain
End Sub

Private Function Translation(ByVal sTriplett As String) As udtAminoAcid

   Dim stiTemp As udtAminoAcid

   If mdiMain.optKindOfOptimize(0) = True Then
      'Optimization for Most GC
      Select Case sTriplett
         Case "GGU", "GGC", "GGA", "GGG"
            stiTemp.sShortcut = "GLY"
            stiTemp.sName = "Glycin"
            stiTemp.sBestTriplett = "GGC"
            stiTemp.sSecondBest = "GGG"
         Case "GCU", "GCC", "GCA", "GCG"
            stiTemp.sShortcut = "ALA"
            stiTemp.sName = "Alanin"
            stiTemp.sBestTriplett = "GCG"
            stiTemp.sSecondBest = "GCC"
         Case "GUU", "GUC", "GUA", "GUG"
            stiTemp.sShortcut = "VAL"
            stiTemp.sName = "Valin"
            stiTemp.sBestTriplett = "GUC"
            stiTemp.sSecondBest = "GUG"
```

```
Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
   stiTemp.sShortcut = "LEU"
   stiTemp.sName = "Leucin"
   stiTemp.sBestTriplett = "CUG"
   stiTemp.sSecondBest = "CUC"
Case "AUA", "AUU", "AUC"
   stiTemp.sShortcut = "ILE"
   stiTemp.sName = "Isoleucin"
   stiTemp.sBestTriplett = "AUC"
Case "UUU", "UUC"
   stiTemp.sShortcut = "PHE"
   stiTemp.sName = "Phenylalanin"
   stiTemp.sBestTriplett = "UUC"
Case "UAU", "UAC"
   stiTemp.sShortcut = "TYR"
   stiTemp.sName = "Tyrosin"
   stiTemp.sBestTriplett = "UAC"
Case "UGG"
   stiTemp.sShortcut = "TRP"
   stiTemp.sName = "Tryptophan"
   stiTemp.sBestTriplett = "UGG"
Case "GAU", "GAC"
   stiTemp.sShortcut = "ASP"
   stiTemp.sName = "Asparaginsäure"
   stiTemp.sBestTriplett = "GAC"
Case "AAU", "AAC"
   stiTemp.sShortcut = "ASN"
   stiTemp.sName = "Asparagin"
   stiTemp.sBestTriplett = "AAC"
Case "GAA", "GAG"
   stiTemp.sShortcut = "GLU"
   stiTemp.sName = "Glutaminsäure"
   stiTemp.sBestTriplett = "GAG"
Case "CAA", "CAG"
   stiTemp.sShortcut = "GLN"
   stiTemp.sName = "Glutamin"
   stiTemp.sBestTriplett = "CAG"
Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
   stiTemp.sShortcut = "SER"
   stiTemp.sName = "Serin"
   stiTemp.sBestTriplett = "AGC"
   stiTemp.sSecondBest = "UCG"
   stiTemp.sThirdBest = "UCC"
Case "ACA", "ACU", "ACG", "ACC"
   stiTemp.sShortcut = "THR"
   stiTemp.sName = "Threonin"
   stiTemp.sBestTriplett = "ACG"
Case "UGU", "UGC"
   stiTemp.sShortcut = "CYS"
   stiTemp.sName = "Cystein"
   stiTemp.sBestTriplett = "UGC"
```

```
        Case "AUG"
          stiTemp.sShortcut = "MET"
          stiTemp.sName = "Methionin"
          stiTemp.sBestTriplett = "AUG"
        Case "CCA", "CCU", "CCG", "CCC"
          stiTemp.sShortcut = "PRO"
          stiTemp.sName = "Prolin"
          stiTemp.sBestTriplett = "CCG"
          stiTemp.sSecondBest = "CCC"
        Case "CAU", "CAC"
          stiTemp.sShortcut = "HIS"
          stiTemp.sName = "Histidin"
          stiTemp.sBestTriplett = "CAC"
        Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
          stiTemp.sShortcut = "ARG"
          stiTemp.sName = "Arginin"
          stiTemp.sBestTriplett = "CGC"
          stiTemp.sSecondBest = "CGG"
        Case "AAA", "AAG"
          stiTemp.sShortcut = "LYS"
          stiTemp.sName = "Lysin"
          stiTemp.sBestTriplett = "AAG"
        Case "UAA", "UAG", "UGA"
          stiTemp.sShortcut = "STP"
          stiTemp.sName = "Stop"
          stiTemp.sBestTriplett = "UGA"
          stiTemp.sSecondBest = "UAG"
      End Select
      Translation = stiTemp
    Else   'Optimization for best frequency
      Select Case sTriplett
        Case "GGU", "GGC", "GGA", "GGG"
          stiTemp.sShortcut = "GLY"
          stiTemp.sName = "Glycin"
          stiTemp.sBestTriplett = "GGC"
          stiTemp.sSecondBest = "GGU"
        Case "GCU", "GCC", "GCA", "GCG"
          stiTemp.sShortcut = "ALA"
          stiTemp.sName = "Alanin"
          stiTemp.sBestTriplett = "GCC"
          stiTemp.sSecondBest = "GCU"
        Case "GUU", "GUC", "GUA", "GUG"
          stiTemp.sShortcut = "VAL"
          stiTemp.sName = "Valin"
          stiTemp.sBestTriplett = "GUG"
          stiTemp.sSecondBest = "GUC"
        Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
          stiTemp.sShortcut = "LEU"
          stiTemp.sName = "Leucin"
          stiTemp.sBestTriplett = "CUG"
          stiTemp.sSecondBest = "CUC"
```

```
Case "AUA", "AUU", "AUC"
   stiTemp.sShortcut = "ILE"
   stiTemp.sName = "Isoleucin"
   stiTemp.sBestTriplett = "AUC"
Case "UUU", "UUC"
   stiTemp.sShortcut = "PHE"
   stiTemp.sName = "Phenylalanin"
   stiTemp.sBestTriplett = "UUC"
Case "UAU", "UAC"
   stiTemp.sShortcut = "TYR"
   stiTemp.sName = "Tyrosin"
   stiTemp.sBestTriplett = "UAC"
Case "UGG"
   stiTemp.sShortcut = "TRP"
   stiTemp.sName = "Tryptophan"
   stiTemp.sBestTriplett = "UGG"
Case "GAU", "GAC"
   stiTemp.sShortcut = "ASP"
   stiTemp.sName = "Asparaginsäure"
   stiTemp.sBestTriplett = "GAC"
Case "AAU", "AAC"
   stiTemp.sShortcut = "ASN"
   stiTemp.sName = "Asparagin"
   stiTemp.sBestTriplett = "AAC"
Case "GAA", "GAG"
   stiTemp.sShortcut = "GLU"
   stiTemp.sName = "Glutaminsäure"
   stiTemp.sBestTriplett = "GAG"
Case "CAA", "CAG"
   stiTemp.sShortcut = "GLN"
   stiTemp.sName = "Glutamin"
   stiTemp.sBestTriplett = "CAG"
Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
   stiTemp.sShortcut = "SER"
   stiTemp.sName = "Serin"
   stiTemp.sBestTriplett = "AGC"
   stiTemp.sSecondBest = "UCC"
   stiTemp.sThirdBest = "UCU"
Case "ACA", "ACU", "ACG", "ACC"
   stiTemp.sShortcut = "THR"
   stiTemp.sName = "Threonin"
   stiTemp.sBestTriplett = "ACC"
Case "UGU", "UGC"
   stiTemp.sShortcut = "CYS"
   stiTemp.sName = "Cystein"
   stiTemp.sBestTriplett = "UGC"
Case "AUG"
   stiTemp.sShortcut = "MET"
   stiTemp.sName = "Methionin"
   stiTemp.sBestTriplett = "AUG"
Case "CCA", "CCU", "CCG", "CCC"
```

```
            stiTemp.sShortcut = "PRO"
            stiTemp.sName = "Prolin"
            stiTemp.sBestTriplett = "CCC"
            stiTemp.sSecondBest = "CCU"
         Case "CAU", "CAC"
            stiTemp.sShortcut = "HIS"
            stiTemp.sName = "Histidin"
            stiTemp.sBestTriplett = "CAC"
         Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
            stiTemp.sShortcut = "ARG"
            stiTemp.sName = "Arginin"
            stiTemp.sBestTriplett = "CGC"
            stiTemp.sSecondBest = "AGG"
         Case "AAA", "AAG"
            stiTemp.sShortcut = "LYS"
            stiTemp.sName = "Lysin"
            stiTemp.sBestTriplett = "AAG"
         Case "UAA", "UAG", "UGA"
            stiTemp.sShortcut = "STP"
            stiTemp.sName = "Stop"
            stiTemp.sBestTriplett = "UGA"
            stiTemp.sSecondBest = "UAG"
      End Select
      Translation = stiTemp
   End If

End Function

Public Function ReverseTranscription() As String
   Dim sTempCode As String
   Dim iLoop As Integer
   For iLoop = 0 To glAminoCounter - 1
      sTempCode = sTempCode + mastrAminoAcids(iLoop).sBestTriplett
   Next
   ReverseTranscription = sTempCode
End Function

Public Sub CountBases(ByVal sTempCode As String)

Dim lLoop As Long
Dim sFormattedCode As String
Dim sActualBase As String
Dim lCodeLength As Long

   lCodeLength = Len(sTempCode)
   For lLoop = 1 To CLng(lCodeLength)
      sActualBase = Mid(sTempCode, lLoop, 1)
      Select Case sActualBase
         Case "A", "a"
            glOptAdenin = glOptAdenin + 1
         Case "U", "u", "T", "t"
```

```
            glOptThymin = glOptThymin + 1
        Case "G", "g"
            glOptGuanin = glOptGuanin + 1
        Case "C", "c"
            glOptCytosin = glOptCytosin + 1
    End Select
  Next lLoop
End Sub
```

Name: basPublics.bas
Code:

```
Attribute VB_Name = "basPublics"
Option Explicit

Public gsOriginalCode As String
Public gsFormattedCode As String
Public glCodeLength As Long

Public glAminoCounter As Long
Public glAdenin As Long
Public glThymin As Long
Public glGuanin As Long
Public glCytosin As Long
Public glOptAdenin As Long
Public glOptThymin As Long
Public glOptGuanin As Long
Public glOptCytosin As Long

Public gbSequenceChanged As Boolean
```

Name: mdiMain.frm
Code:

```
VERSION 5.00
Object = "{38911DA0-E448-11D0-84A3-00DD01104159}#1.1#0"; "COMCT332.OCX"
Begin VB.MDIForm mdiMain
  BackColor      =  &H00FFFFFF&
  Caption        =  "Curevac_RNA_Analyzer"
  ClientHeight   =  8745
  ClientLeft     =  165
  ClientTop      =  735
  ClientWidth    =  12255
  Icon           =  "mdiMain.frx":0000
  LinkTopic      =  "MDIForm1"
  Picture        =  "mdiMain.frx":058A
  StartUpPosition =  3  'Windows Default
  WindowState    =  2  'Maximized
  Begin ComCtl3.CoolBar cbarMain
```

```
Align           =  1  'Align Top
Height          =  885
Left            =  0
TabIndex        =  0
Top             =  0
Width           =  12255
_ExtentX        =  21616
_ExtentY        =  1561
BandCount       =  1
BandBorders     =  0  'False
_CBWidth        =  12255
_CBHeight       =  885
_Version        =  "6.7.8988"
MinHeight1      =  825
Width1          =  4995
FixedBackground1=  0  'False
NewRow1         =  0  'False
Begin VB.OptionButton optKindOfOptimize
  Caption       =  "Best Frequency"
  Height        =  255
  Index         =  1
  Left          =  4920
  TabIndex      =  7
  Top           =  480
  Value         =  -1  'True
  Width         =  1575
End
Begin VB.OptionButton optKindOfOptimize
  Caption       =  "Best GC"
  Height        =  255
  Index         =  0
  Left          =  4920
  TabIndex      =  6
  Top           =  120
  Width         =  1575
End
Begin VB.CommandButton cmdShowInput
  Caption       =  "Input"
  Height        =  765
  Left          =  0
  Picture       =  "mdiMain.frx":0C2A
  Style         =  1  'Graphical
  TabIndex      =  5
  Top           =  0
  Width         =  840
End
Begin VB.CommandButton cmdShowOptions
  Caption       =  "Options"
  Enabled       =  0  'False
  Height        =  760
  Left          =  3480
```

```
    Picture         =  "mdiMain.frx":0F34
    Style           =  1  'Graphical
    TabIndex        =  4
    Top             =  0
    Width           =  735
  End
  Begin VB.CommandButton cmdShowStatistics
    Caption         =  "Statistics"
    Enabled         =  0  'False
    Height          =  760
    Left            =  2640
    Picture         =  "mdiMain.frx":131E
    Style           =  1  'Graphical
    TabIndex        =  3
    Top             =  0
    Width           =  735
  End
  Begin VB.CommandButton cmdShowDisplay
    Caption         =  "Display"
    Enabled         =  0  'False
    Height          =  760
    Left            =  1800
    Picture         =  "mdiMain.frx":1628
    Style           =  1  'Graphical
    TabIndex        =  2
    Top             =  0
    Width           =  735
  End
  Begin VB.CommandButton cmdShowOutput
    Caption         =  "Output"
    Enabled         =  0  'False
    Height          =  760
    Left            =  960
    Picture         =  "mdiMain.frx":20E2
    Style           =  1  'Graphical
    TabIndex        =  1
    Top             =  0
    Width           =  735
  End
End
Begin VB.Menu mnuMain
  Caption           =  "&Input..."
  Index             =  0
End
Begin VB.Menu mnuMain
  Caption           =  "&Results"
  Index             =  1
  Begin VB.Menu mnuResults
    Caption         =  "&Output..."
    Enabled         =  0  'False
    Index           =  0
```

```
  End
  Begin VB.Menu mnuResults
    Caption        =  "&Display..."
    Enabled        =  0  'False
    Index          =  1
  End
  Begin VB.Menu mnuResults
    Caption        =  "&Statistics..."
    Enabled        =  0  'False
    Index          =  2
  End
End
Begin VB.Menu mnuMain
  Caption          =  "E&xtras"
  Index            =  4
  Begin VB.Menu mnuExtras
    Caption        =  "&Language"
    Index          =  0
    Begin VB.Menu mnuLanguage
      Caption      =  "English"
      Checked      =  -1  'True
      Index        =  0
    End
    Begin VB.Menu mnuLanguage
      Caption      =  "&German"
      Enabled      =  0  'False
      Index        =  1
    End
    Begin VB.Menu mnuLanguage
      Caption      =  "&French"
      Enabled      =  0  'False
      Index        =  2
    End
  End
  Begin VB.Menu mnuExtras
    Caption        =  "&Options..."
    Enabled        =  0  'False
    Index          =  1
  End
  Begin VB.Menu mnuExtras
    Caption        =  "-"
    Index          =  2
  End
  Begin VB.Menu mnuExtras
    Caption        =  "&About"
    Index          =  3
  End
End
Begin VB.Menu mnuMain
  Caption          =  "&Windows"
  Index            =  5
```

```
    WindowList      =   -1  'True
    Begin VB.Menu mnuWindows
      Caption       =   "Tile &Horizontally"
      Index         =   0
    End
    Begin VB.Menu mnuWindows
      Caption       =   "Tile &Vertically"
      Index         =   1
    End
    Begin VB.Menu mnuWindows
      Caption       =   "&Cascade"
      Index         =   2
    End
  End
  Begin VB.Menu mnuMain
    Caption         =   "&Exit"
    Index           =   6
  End
End
Attribute VB_Name = "mdiMain"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub cmdShowDisplay_Click()
  Call frmDisplay.Show
End Sub

Private Sub cmdShowInput_Click()
  Call frmInput.Show
End Sub

Private Sub cmdShowOptions_Click()
  Call frmOptions.Show
End Sub

Private Sub cmdShowOutput_Click()
  Call frmOutPut.Show
End Sub

Private Sub cmdShowStatistics_Click()
  Call frmStatistics.Show
End Sub

Private Sub MDIForm_Load()
  Call InitControls
  Call frmInput.Show
End Sub
```

```
Private Sub InitControls()

Dim lNextLeft As Long
Const lSpace As Long = 60
Const lButtonWidth As Long = 850
Const lButtonHeight As Long = 770

   lNextLeft = lSpace + lButtonWidth

   With cmdShowInput
      .Top = lSpace
      .Left = lSpace
      .Width = lButtonWidth
      .Height = lButtonHeight
   End With

   With cmdShowOutput
      .Top = lSpace
      .Left = lNextLeft + lSpace
      .Width = lButtonWidth
      .Height = lButtonHeight
   End With

   With cmdShowDisplay
      .Top = lSpace
      .Left = lNextLeft * 2 + lSpace
      .Width = lButtonWidth
      .Height = lButtonHeight
   End With

   With cmdShowStatistics
      .Top = lSpace
      .Left = lNextLeft * 3 + lSpace
      .Width = lButtonWidth
      .Height = lButtonHeight
   End With

   With cmdShowOptions
      .Top = lSpace
      .Left = lNextLeft * 4 + lSpace
      .Width = lButtonWidth
      .Height = lButtonHeight
   End With

   cbarMain.Height = lButtonHeight + 2 * lSpace


End Sub

Private Sub mnuMain_Click(Index As Integer)
```

```
    Select Case Index
        Case 0
            Call frmInput.Show
        Case 6
            Unload Me
    End Select

End Sub

Private Sub mnuResults_Click(Index As Integer)

    Select Case Index
        Case 0 'Output
            Call frmOutPut.Show
        Case 1
            Call frmDisplay.Show
        Case 2 'Statistische Auswertung
'           Call frmStatistics.Show
    End Select

End Sub

Private Sub mnuExtras_Click(Index As Integer)

    Select Case Index
        Case 0 'Output
            '
        Case 1
            Call frmDisplay.Show
        Case 3
            Call frmAbout.Show(vbModal)
    End Select

End Sub


Private Sub mnuWindows_Click(Index As Integer)

    Select Case Index
        Case 0 'Horizontal
            Me.Arrange (vbTileHorizontal)
        Case 1 'Vertikal
            Me.Arrange (vbTileVertical)
        Case 2 'Kaskadieren
            Me.Arrange (vbCascade)
    End Select

End Sub
```

Name: frmInput.frm

Code:

```
VERSION 5.00
Object = "{F9043C88-F6F2-101A-A3C9-08002B2F49FB}#1.2#0"; "COMDLG32.OCX"
Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}#1.2#0"; "RICHTX32.OCX"
Begin VB.Form frmInput
  Caption        =  "Input"
  ClientHeight   =  5730
  ClientLeft     =  60
  ClientTop      =  345
  ClientWidth    =  13200
  Icon           =  "frmInput.frx":0000
  LinkTopic      =  "Form1"
  MDIChild       =  -1  'True
  ScaleHeight    =  5730
  ScaleWidth     =  13200
  WindowState    =  2  'Maximized
  Begin RichTextLib.RichTextBox txtFormatted
    Height         =  1815
    Left           =  120
    TabIndex       =  3
    Top            =  360
    Width          =  12735
    _ExtentX       =  22463
    _ExtentY       =  3201
    _Version       =  393217
    BorderStyle    =  0
    Enabled        =  -1  'True
    ScrollBars     =  2
    DisableNoScroll =  -1  'True
    TextRTF        =  $"frmInput.frx":030A
    BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}
      Name           =  "Fixedsys"
      Size           =  9
      Charset        =  0
      Weight         =  400
      Underline      =  0  'False
      Italic         =  0  'False
      Strikethrough  =  0  'False
    EndProperty
  End
  Begin VB.OptionButton optSequence
    Caption        =  "Formatted"
    Height         =  255
    Index          =  1
    Left           =  2760
    Style          =  1  'Graphical
    TabIndex       =  2
    Top            =  0
    Value          =  -1  'True
    Width          =  855
  End
```

```
  Begin VB.OptionButton optSequence
    Caption         =   "Original"
    Height          =   255
    Index           =   0
    Left            =   1920
    Style           =   1  'Graphical
    TabIndex        =   1
    Top             =   0
    Width           =   855
  End
  Begin VB.CommandButton cmdLoad
    Caption         =   "Load Sequence"
    Height          =   285
    Left            =   0
    TabIndex        =   0
    Top             =   0
    Width           =   1695
  End
  Begin MSComDlg.CommonDialog CommonDialog1
    Left            =   3840
    Top             =   -120
    _ExtentX        =   847
    _ExtentY        =   847
    _Version        =   393216
  End
End
Attribute VB_Name = "frmInput"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit


Private Sub Form_Activate()
   Me.ZOrder
   Call InitControls
End Sub

Public Sub LoadSequence()

Dim sFile As String
Dim sTempCode As String

   On Error GoTo Errorhandler
   With CommonDialog1
      .FileName = App.Path & "\SampleRNA.txt"
      .CancelError = True
      Call .ShowOpen
   End With
   If Len(CommonDialog1.FileName) > 0 Then
```

```
      gbSequenceChanged = True
      gsFormattedCode = ""
      gsOriginalCode = ""
      sFile = CommonDialog1.FileName
      Open sFile For Input As #1
      Input #1, gsOriginalCode
      Close #1
      Call FormatText
      Call FillSequence

      mdiMain.cmdShowDisplay.Enabled = True
      mdiMain.cmdShowOutput.Enabled = True
      mdiMain.cmdShowStatistics.Enabled = True
   End If

Exit Sub

Errorhandler:

Dim lAnswer As Long

   If Err.Number <> cdlCancel Then
      lAnswer = MsgBox("Datei konnte nicht geladen werden", vbRetryCancel,
"Dateiproblem")
      If lAnswer = vbRetry Then
         Resume
      End If
   End If

End Sub

Private Sub FormatText()

Dim iLoop As Integer
Dim sActualBase As String

   glCodeLength = Len(gsOriginalCode)
   For iLoop = 0 To glCodeLength - 1
      sActualBase = Mid(gsOriginalCode, iLoop + 1, 1)
      If iLoop Mod 3 = 0 Then
         gsFormattedCode = gsFormattedCode + " "
         glAminoCounter = glAminoCounter + 1
      End If
      Select Case sActualBase
         Case "A", "a"
            gsFormattedCode = gsFormattedCode + "A"
            glAdenin = glAdenin + 1
         Case "U", "u", "T", "t"
            gsFormattedCode = gsFormattedCode + "U"
            glThymin = glThymin + 1
         Case "G", "g"
```

```
            gsFormattedCode = gsFormattedCode + "G"
            glGuanin = glGuanin + 1
        Case "C", "c"
            gsFormattedCode = gsFormattedCode + "C"
            glCytosin = glCytosin + 1
    End Select
  Next iLoop
  gsFormattedCode = Trim(gsFormattedCode)

End Sub

Private Sub FillSequence()
  If optSequence(0).Value = True Then
    txtFormatted.Text = gsOriginalCode
  Else
    txtFormatted.Text = gsFormattedCode
  End If

End Sub

Private Sub cmdShowOutput_Click()
  Call frmOutPut.Show
End Sub

Private Sub cmdLoad_Click()
  Call LoadSequence
End Sub

Private Sub InitControls()

Const lSpace As Long = 40
Const lButtonHeight As Long = 285

  With cmdLoad
    .Top = lSpace
    .Height = lButtonHeight
    .Left = lSpace
  End With

  With optSequence(0)
    .Top = lSpace
    .Height = lButtonHeight
    .Left = cmdLoad.Left + cmdLoad.Width + lSpace + 200
  End With

  With optSequence(1)
    .Top = lSpace
    .Height = lButtonHeight
    .Left = optSequence(0).Left + optSequence(0).Width
  End With
```

```
    With txtFormatted
        .Top = cmdLoad.Height + 2 * lSpace
        .Left = lSpace
        .Width = Me.ScaleWidth
        .Height = Me.ScaleHeight - txtFormatted.Top
    End With
End Sub

Private Sub optSequence_Click(Index As Integer)
    Call FillSequence
End Sub
```

Name: frmOutPut.frm
Code:

```
VERSION 5.00
Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}#1.2#0"; "RICHTX32.OCX"
Begin VB.Form frmOutPut
   Caption        =   "OutPut"
   ClientHeight   =   9120
   ClientLeft     =   60
   ClientTop      =   345
   ClientWidth    =   10215
   Icon           =   "frmOutPut.frx":0000
   LinkTopic      =   "Form1"
   MDIChild       =   -1  'True
   ScaleHeight    =   9120
   ScaleWidth     =   10215
   Begin RichTextLib.RichTextBox txtOptimized
      Height         =   2655
      Left           =   0
      TabIndex       =   0
      Top            =   480
      Width          =   9855
      _ExtentX       =   17383
      _ExtentY       =   4683
      _Version       =   393217
      Enabled        =   -1  'True
      ScrollBars     =   2
      DisableNoScroll =   -1  'True
      TextRTF        =   $"frmOutPut.frx":0442
      BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}
         Name           =   "Fixedsys"
         Size           =   9
         Charset        =   0
         Weight         =   400
         Underline      =   0   'False
         Italic         =   0   'False
         Strikethrough  =   0   'False
      EndProperty
```

```
  End
End
Attribute VB_Name = "frmOutPut"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub Form_Activate()

   Me.ZOrder
   Call InitControls
   If gbSequenceChanged = True Then
      Call BuildAminoChain(gsFormattedCode, False)
   End If
   gbSequenceChanged = False
   txtOptimized.Text = LCase(ReverseTranscription)

End Sub


Private Sub InitControls()

Const lSpace As Long = 40

   With txtOptimized
      .Top = lSpace
      .Left = lSpace
      .Width = Me.ScaleWidth + lSpace
      .Height = Me.ScaleHeight + lSpace
   End With
End Sub
```

Name: frmDisplay.frm
Code:

```
VERSION 5.00
Object = "*\ACurevac_Genetic_Controls.vbp"
Begin VB.Form frmDisplay
   Caption        =   "Display"
   ClientHeight   =   8205
   ClientLeft     =   60
   ClientTop      =   345
   ClientWidth    =   8745
   Icon           =   "frmDisplay.frx":0000
   LinkTopic      =   "Form1"
   MDIChild       =   -1  'True
   ScaleHeight    =   8205
   ScaleWidth     =   8745
```

```
  ShowInTaskbar   =  0  'False
  Begin VB.OptionButton optNucleotids
    Caption       =  "Original"
    Height        =  255
    Index         =  0
    Left          =  0
    Style         =  1  'Graphical
    TabIndex      =  1
    Top           =  0
    Width         =  855
  End
  Begin VB.OptionButton optNucleotids
    Caption       =  "Optimized"
    Height        =  255
    Index         =  1
    Left          =  840
    Style         =  1  'Graphical
    TabIndex      =  0
    Top           =  0
    Value         =  -1  'True
    Width         =  855
  End
  Begin Curevac_Genetic_Controls.Curevac_Amino Curevac_Amino1
    Height        =  495
    Index         =  0
    Left          =  240
    Top           =  360
    Visible       =  0  'False
    Width         =  735
    _ExtentX      =  1296
    _ExtentY      =  873
  End
End
Attribute VB_Name = "frmDisplay"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub Form_Activate()

   Me.ZOrder
   If gbSequenceChanged = True Then
      Call BuildAminoChain(gsFormattedCode, True)
   End If
   gbSequenceChanged = False

End Sub
```

```
Private Sub optNucleotids_Click(Index As Integer)

Dim bShowOriginal As Boolean

   If Index = 0 Then 'Original Clicked
      bShowOriginal = True
   Else 'Optimized Clicked
      bShowOriginal = False
   End If
   Call RebuildChain(bShowOriginal)

End Sub
```

Name: frmStatistics.frm
Code:

```
VERSION 5.00
Begin VB.Form frmStatistics
  Caption        =  "Statistics"
  ClientHeight   =  6450
  ClientLeft     =  60
  ClientTop      =  345
  ClientWidth    =  8595
  Icon           =  "frmStatistics.frx":0000
  LinkTopic      =  "Form1"
  MDIChild       =  -1  'True
  ScaleHeight    =  6450
  ScaleWidth     =  8595
  Begin VB.Frame Frame1
    Caption      =  "original"
    Height       =  2655
    Left         =  120
    TabIndex     =  9
    Top          =  720
    Width        =  2175
    Begin VB.Label lblAdenin
      BorderStyle    =  1  'Fixed Single
      Height         =  495
      Left           =  840
      TabIndex       =  17
      Top            =  240
      Width          =  1215
    End
    Begin VB.Label lblCytosin
      BorderStyle    =  1  'Fixed Single
      Height         =  495
      Left           =  840
      TabIndex       =  16
      Top            =  2040
      Width          =  1215
```

```
  End
  Begin VB.Label lblGuanin
    BorderStyle     =  1  'Fixed Single
    Height          =  495
    Left            =  840
    TabIndex        =  15
    Top             =  1440
    Width           =  1215
  End
  Begin VB.Label lblThymin
    BorderStyle     =  1  'Fixed Single
    Height          =  495
    Left            =  840
    TabIndex        =  14
    Top             =  840
    Width           =  1215
  End
  Begin VB.Label Label14
    Caption         =  "Cytosin"
    Height          =  255
    Left            =  120
    TabIndex        =  13
    Top             =  2160
    Width           =  615
  End
  Begin VB.Label Label2
    Caption         =  "Thymin"
    Height          =  255
    Left            =  120
    TabIndex        =  12
    Top             =  960
    Width           =  615
  End
  Begin VB.Label Label3
    Caption         =  "Guanin"
    Height          =  255
    Left            =  120
    TabIndex        =  11
    Top             =  1560
    Width           =  615
  End
  Begin VB.Label Label1
    Caption         =  "Adenin"
    Height          =  255
    Left            =  120
    TabIndex        =  10
    Top             =  360
    Width           =  615
  End
End
Begin VB.Frame Frame2
```

```
Caption      = "optimized"
Height       = 2655
Left         = 2520
TabIndex     = 0
Top          = 720
Width        = 2175
Begin VB.Label Label4
  Caption    = "Adenin"
  Height     = 255
  Left       = 120
  TabIndex   = 8
  Top        = 360
  Width      = 615
End
Begin VB.Label Label6
  Caption    = "Cytosin"
  Height     = 255
  Left       = 120
  TabIndex   = 7
  Top        = 2160
  Width      = 615
End
Begin VB.Label Label8
  Caption    = "Guanin"
  Height     = 255
  Left       = 120
  TabIndex   = 6
  Top        = 1560
  Width      = 615
End
Begin VB.Label Label12
  Caption    = "Thymin"
  Height     = 255
  Left       = 120
  TabIndex   = 5
  Top        = 960
  Width      = 615
End
Begin VB.Label lblOptAdenin
  BorderStyle  = 1 'Fixed Single
  Height     = 495
  Left       = 840
  TabIndex   = 4
  Top        = 240
  Width      = 1215
End
Begin VB.Label lblOptThymin
  BorderStyle  = 1 'Fixed Single
  Height     = 495
  Left       = 840
  TabIndex   = 3
```

```
      Top             =   840
      Width           =   1215
    End
    Begin VB.Label lblOptGuanin
      BorderStyle     =   1  'Fixed Single
      Height          =   495
      Left            =   840
      TabIndex        =   2
      Top             =   1440
      Width           =   1215
    End
    Begin VB.Label lblOptCytosin
      BorderStyle     =   1  'Fixed Single
      Height          =   495
      Left            =   840
      TabIndex        =   1
      Top             =   2040
      Width           =   1215
    End
  End
  Begin VB.Label lblBases
    BorderStyle     =   1  'Fixed Single
    Height          =   495
    Left            =   1320
    TabIndex        =   19
    Top             =   120
    Width           =   1215
  End
  Begin VB.Label Label5
    Caption         =   "Sum of bases"
    Height          =   255
    Left            =   120
    TabIndex        =   18
    Top             =   240
    Width           =   1095
  End
End
Attribute VB_Name = "frmStatistics"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub Form_Activate()

   Me.ZOrder

   lblBases.Caption = CStr(glAminoCounter * 3)
   lblAdenin.Caption = CStr(glAdenin)
   lblThymin.Caption = CStr(glThymin)
```

```
    lblGuanin.Caption = CStr(glGuanin)
    lblCytosin.Caption = CStr(glCytosin)

    Call CountBases(frmOutPut.txtOptimized.Text)
    lblOptAdenin.Caption = CStr(glOptAdenin)
    lblOptThymin.Caption = CStr(glOptThymin)
    lblOptGuanin.Caption = CStr(glOptGuanin)
    lblOptCytosin.Caption = CStr(glOptCytosin)

End Sub
```

Name: frmOptions.frm
Code:

```
VERSION 5.00
Begin VB.Form frmOptions
   Caption        =   "Options"
   ClientHeight   =   3915
   ClientLeft     =   60
   ClientTop      =   345
   ClientWidth    =   5760
   Icon           =   "frmOptions.frx":0000
   LinkTopic      =   "Form1"
   MDIChild       =   -1  'True
   ScaleHeight    =   3915
   ScaleWidth     =   5760
End
Attribute VB_Name = "frmOptions"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit
```

Name: frmAbout.frm
Code:

```
VERSION 5.00
Begin VB.Form frmAbout
   BorderStyle    =   3  'Fixed Dialog
   Caption        =   "About Curevac RNA-Analyzer"
   ClientHeight   =   2490
   ClientLeft     =   2340
   ClientTop      =   1935
   ClientWidth    =   6195
   ClipControls   =   0   'False
   LinkTopic      =   "Form2"
   MaxButton      =   0   'False
   MinButton      =   0   'False
```

```
ScaleHeight    =  1718.642
ScaleMode      =  0  'User
ScaleWidth     =  5817.425
ShowInTaskbar  =  0  'False
Begin VB.PictureBox picIcon
  AutoSize       =  -1  'True
  ClipControls   =  0  'False
  Height         =  1215
  Left           =  120
  Picture        =  "frmAbout.frx":0000
  ScaleHeight    =  811.195
  ScaleMode      =  0  'User
  ScaleWidth     =  2401.98
  TabIndex       =  1
  Top            =  120
  Width          =  3480
End
Begin VB.CommandButton cmdOK
  Cancel         =  -1  'True
  Caption        =  "OK"
  Default        =  -1  'True
  Height         =  345
  Left           =  2520
  TabIndex       =  0
  Top            =  2040
  Width          =  1260
End
Begin VB.Label lblCompany
  Caption        =  "lblCompany"
  Height         =  255
  Left           =  3720
  TabIndex       =  5
  Top            =  1200
  Width          =  2295
End
Begin VB.Line Line1
  BorderColor    =  &H00808080&
  BorderStyle    =  6  'Inside Solid
  Index          =  1
  X1             =  112.686
  X2             =  5746.996
  Y1             =  1325.218
  Y2             =  1325.218
End
Begin VB.Label lblDescription
  Caption        =  "App Description"
  ForeColor      =  &H00000000&
  Height         =  330
  Left           =  120
  TabIndex       =  2
  Top            =  1560
```

```
    Width          =   5925
  End
  Begin VB.Label lblTitle
    Caption        =   "Application Title"
    ForeColor      =   &H00000000&
    Height         =   360
    Left           =   3720
    TabIndex       =   3
    Top            =   240
    Width          =   2325
  End
  Begin VB.Line Line1
    BorderColor    =   &H00FFFFFF&
    BorderWidth    =   2
    Index          =   0
    X1             =   112.686
    X2             =   5746.996
    Y1             =   1325.218
    Y2             =   1325.218
  End
  Begin VB.Label lblVersion
    Caption        =   "Version"
    Height         =   225
    Left           =   3720
    TabIndex       =   4
    Top            =   720
    Width          =   2325
  End
End
Attribute VB_Name = "frmAbout"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub cmdOK_Click()
  Unload Me
End Sub


Private Sub Form_Load()
   Me.Caption = "About " & App.Title
   lblVersion.Caption = "Version " & App.Major & "." & App.Minor & "." & App.Revision
   lblTitle.Caption = App.Title
   lblDescription.Caption = App.FileDescription
   lblCompany.Caption = App.CompanyName
End Sub
```